

28 October 1982

Cable in haste, repent at leisure

The British government is exaggerating the benefits of cable television and underestimating the cost. It should urgently investigate the economics of its proposals.

Is a new technology a way of transforming society or merely of inducing people to spend more? These questions are not raised in the advice given recently to the British government by the committee under Lord Hunt, former Secretary to the Cabinet. The Hunt committee, given only six months for its task, dutifully told the government just what it wanted to know: how to provide safeguards for national broadcasting while encouraging the expansion of a national cable system with at least 30 video channels and an interactive facility.

The committee's recommendations would leave cable, by broadcasting standards, virtually free of regulation. Hunt did offer some sensible proposals for allocating franchises without getting into the morass of local government corruption and greed that is slowing down cable's progress in the United States. Briefly, he would all but exclude local government from the franchising process, leaving that to a national cable authority that should itself subsequently supervise cable operators very lightly, with no day to day supervision of programmes or even of fees — Hunt has great faith in market forces — but with the risk that their eight-year franchises will not be renewed.

Debate

Lord Hunt has asked that a debate be begun. This debate is doomed to be a disappointment. Cable is misunderstood and misinterpreted, by those like the broadcasting institutions who feel their survival threatened, and by the government which sees "rewiring Britain" as a handy slogan for the next election to use against reminders of three million unemployed. The British Broadcasting Corporation and the Independent Broadcasting Authority are proclaiming the imminent end of the British tradition of balanced, public service broadcasting, without appreciating that cable television is, by its multi-channelled nature, complementary to and not competitive with national mass-oriented airborne networks. Other custodians of public taste, elected, self-appointed and both, will be up in arms about the prospect that cable television subscribers will be able more cheaply than with their video recorders to watch soft porn. A few judicious compromises in the next few months should give the British government an easy ride through the House of Commons some time next summer for the bill that it will need to let franchise holders dig up the streets, and further to erode the old monopoly powers of British Telecom, as well as the power of local governments to grant wayleaves. The danger is that the debate that will occupy the next few months will disastrously neglect the questions that should be asked — is the British government right in thinking that cable television by itself will bring prosperity, or is it simply economic candy floss?

The government is in no doubt. Under the guidance of its full-time cheerleader, Mr Kenneth Baker, once with the software house Logica but now a minister of state at the Department of Industry, the British government has slipped into the habit of believing that Britain is backward in manufacturing electronic equipment that ordinary people buy (which is true) and must therefore embrace every opportunity to change its ways (which is not true). This is the spirit in which it nominated 1982 as "Information Technology Year" — and in which it has begun complaining that people have not taken it seriously enough. The danger now is that the government may be entirely mistaken in

believing what at least Mr Baker has persuaded himself — that cable television equals prosperity.

American experience shows that cannot be true. If the government can wave its legislative green flag next summer, conveniently ahead of the next general election, there will be plenty of people willing to invest in companies scrambled together to apply for franchises to operate cable systems. At this stage it is impossible to estimate what the total capital cost will be. Much will depend on the technical standards that will not even have been drafted until the spring, and much too on other requirements that may be laid in the path of cabling around Britain which make entrepreneurs judge that the venture will not be worth the risk. The government's (or rather Mr Baker's) own guess is £3,000 million, spread perhaps over five years. By the yardstick of the present level of private investment in British industry and commerce (as distinct from lending to the government) this is small beer — perhaps five per cent of the total over the next five years. But is this the way that a government worried stiff about the survival of manufacturing industry would wish to see industrial investment channelled in the next few years? On the face of things, for a country dependent to the tune of 30 per cent of its Gross National Product on exports, industries capable of making known electronics products more competitively would be better bets. Naturally, if the British government is clever enough to insist in its approval of technical standards that British cable systems are technically just the right distance ahead of comparable systems elsewhere, there could be a short spell in which British manufacturers of transmission and terminal equipment enjoyed an advantage over their competitors. The unhappy snag is that the specification of exacting standards will frighten off investors and their potential subscribers, the danger in the Concorde syndrome (too much too soon). Sceptical taxpayers should have in mind a realistic lower limit of what to expect — the chance that a suddenly increased demand for television signals to put out on cable systems will give more British broadcasters, who appear to have some skill, a greater incentive to sell overseas.

Unemployment

But will not a new investment project on this scale do something to lift the British economy out of its doldrums? For you cannot spend £3,000 million (or, more probably, twice as much) without creating jobs. And the money spent will finish up in the pockets of those concerned, not be buried in trenches in the ground. That seems to be an important part of the government's calculation, but it raises two questions, one of arithmetic and one of principle. At present wage rates, a new investment of even £6,000 million would pay for only one million man-years of work, but since half the investment would probably finish up with the manufacturers of equipment now working well below capacity (and able to increase that only by buying more raw materials from abroad), a five-year programme is unlikely to provide more than 100,000 new jobs. The arithmetical question is whether a special programme to "cable Britain" can be justified by a mere three per cent (or less) reduction of the dole queues. The question of principle is why, if the government is as eager as it seems to give cable television a green light, it should steadfastly have resisted the pleadings of the Confederation of British Industry and of the Trades Union Congress that monetary restraint should be relaxed



(to the tune of £2,000 million and twice as much respectively). The obvious monetarist rejoinder, that it matters whether the money comes from private pockets or from the public revenue, does not cover the objection that the encouragement of cable television, essentially a domestic service industry, will ironically make it harder for manufacturing industry to lay its hands on investment capital just at the time when the government's own demands on the capital markets are at last abating.

None of this implies that the British government should not do the decent thing and fall in with proposals like Hunt's for regulating the development of cable television in Britain, but the case for following such a course stems not from economics but from equity — entrepreneurs wishing to invest in cable systems should be free to do so, and subscribers wishing to subscribe should equally be free to pay their money. For would-be cable operators (and their accountants) the crucial question will be how much of that money there will be. In Britain, the demand curve (what proportion at which price per month) has not yet been drawn. For the government, however, it is more important to know where subscribers' monthly payments to the still non-existent cable television operators' services will come from. Other things being equal, the economists' loose translation of *ceteris paribus*, subscribers will be able to pay their monthly bills only by saving less (which would be bad for the rest of British industry) or by consuming less, which would be good for British industry only if, against recent form, British consumers discriminated against imported goods. In the acrimony about the government's intentions that will occupy the next few months, Mr Baker's little experiment with cable should thus be clearly seen for what it is: not a magic programme for the reinvigoration of British industry but a step in the right (because inevitable) direction taken too fast.

No future for the Greens

The Hamburg elections may show that extremist single-issue parties will not last long.

Ironically, the British would-be equivalent of the West German Greens, a political party dedicated to environmental causes, held its first public meeting last weekend at Bridlington (a small windy coastal town in Yorkshire) while the Greens of Hamburg were reluctantly voting for a new general election in the *Land*. The British party, still in its infancy but conscious that it has to be organized to fight elections within a year or so, is full of dreams but unsure how to make them come true. The West German party, on the other hand, having made a principle of its policy not to form coalitions with conventional political parties, and conscious that nine months of impotence in the Hamburg parliament have weakened its reputation, can only look forward to the next Hamburg election with trepidation. This time, the voters are likely to take the view that politics is too serious to be left to environmentalists. The Hamburg Greens are unlikely to repeat their spectacular success earlier this year.

The moral should be more widely understood than it has been in Hamburg or in Bridlington. Environmental issues are socially important and politically unavoidable. Air pollution standards, or public policies on nuclear power, concern (or should concern) voters, and political parties should know where they stand on them. No doubt from frustration at the deafness of the established parties, the West German Greens have gone off on their own, winning votes by exaggeration but for that reason unable to exercise the power thus derived. They would have been better advised to have pressed more moderate versions of their arguments on the existing parties. The same is true of those who met at Bridlington at the weekend.

American contributors please note

Authors in the United States and Canada are reminded that they should send their manuscripts either direct to London or to the Washington office, whose address is *Nature*, 991 National Press Building, Washington, DC 20045; manuscripts should *not* be sent to New York.

Graduate student choice

The British research councils may yet see the value of manpower planning.

Armies have cannon-fodder, political parties have lobby-fodder and research laboratories have graduate students. This widespread assumption is nowhere more faithfully endorsed than in Britain, where the chief sources of support for graduate education, three of the five research councils, all make a practice of delegating to university departments the responsibility for deciding which graduating students should hold the publicly-supported studentships by which most British graduate students try to keep alive. The standard procedure is that selected university departments will be given a quota of studentships which they may offer (within limits) to would-be graduate students.

Naturally, university and polytechnic departments prize their quotas as jealously as they would a kind of collective Nobel medal. For a quota is at once a guarantee that research will be sustained at some level and a powerful lever in the endless interdepartmental competition for resources. This is one reason why Sir Keith Joseph, the British Secretary of State for Education and Science whom everybody hates to love, will have made more enemies by endorsing last week the notion that the Social Science Research Council should deliberately move away from its quota system. Sir Keith should, however, have gone further.

The proposal that the Social Science Research Council should make fewer awards on the quota system stems from Lord Rothschild's report on the working of the council (see *Nature* 27 May, p.254) and was intended as a way of reconciling the need for some kind of manpower planning with sympathy for small university departments with nil quotas in which, the social sciences being what they are, research can never get off the ground. The Rothschild proposal was that quotas should attach to subjects, not to university departments. Sir Keith wants smaller quotas and more of the available awards put into a pool for competition by all university departments. Why not? And why should the same rule not apply to the Science and Engineering Research Council, which makes 3,250 or so new awards each year to graduate students in science and engineering?

In the face of the inevitable and fierce arguments that no change is best, the case for loosening the present quota system deserves a wider hearing. First, with the British university system now in unprecedented flux, the committees that decide quota allocations can be sure only that these will not match the needs of students and the capabilities of departments. Second, the quota system has the effect of spreading graduate students too widely for the interests of graduate education (and the inclinations of the students): a handful of graduate students in a small department does not constitute a graduate school, however much it may flatter the egos of the department heads involved. Third, with even the best research departments under continuing financial pressure, there are good reasons why the best of them should be given a better chance of survival by competition for some share of a pool. The research councils are always saying that they want to concentrate resources where they can be used effectively. Has not the time come for them to back words with deeds, and money?

But competition for and among would-be graduate students is surely a recipe for anarchy. That is what British university departments used to the quota system will be saying. They will of course be wrong. How else, it must be asked, do the outstanding graduate schools in the United States recruit their students? And there is even a British experiment along these lines — the apparently successful award this summer of a proportion of the graduate studentships available from the Science and Engineering Research Council for graduate students in mathematics. When will the others have the courage to follow suit? If mathematicians have been able to put graduate students up for grabs among them without bringing research to a halt, surely physicists, biochemists and others could take the risk. Equity among universities (rewarding quality) and students requires no less.

Belgian universities hit trouble

Money problem complicated by language split

Brussels

The beginning to the new academic year has been depressing for Belgian academics. Almost all of Belgium's oldest and most renowned universities are having to make sweeping staff cutbacks and to reduce the amount of research work possible for the next seven years. Under the special powers awarded to the present centre-right coalition government at the beginning of the year, the government has been free to take any measures it sees fit to stem Belgium's rising public spending deficit. The universities, at least those running large deficits, were given an ultimatum either to put their own financial houses in order, or to have the government intervening directly.

Anxious to preserve their autonomy, the university authorities have been sharpening the axe themselves and their plans are now being studied by the central government. Defiant talk of forming a united front against the government's demands, despite strikes and occupation of university buildings at the Catholic University of Louvain, has subsided into a grumbling acquiescence.

The truth is that the universities are divided and find themselves in widely differing situations. With the important exception of Ghent, only the Walloon or French-speaking universities are faced with a problem. The Flemish universities are by and large the creation of the 1960s, with a younger staff and a more streamlined and efficient administration.

The older francophone universities, which began to run into trouble towards the end of the 1970s, now have their backs against the wall. According to Hervé Hasquin, rector of the Université Libre de Bruxelles (ULB), his university has seen its budget dwindle by 25 per cent from as far back as 1975.

This year alone, the measures taken by Education Minister Michel Tromont have slashed ULB's budget by BF 300 million (\$6 million) and in the country as a whole savings of around BF 2,000 million (\$40 million) will probably have been achieved by the end of the year.

The government has reduced the payments per student, by which the universities finance themselves, by 2.5 per cent and has abolished grants for students on postgraduate courses. As in the United Kingdom, foreign students have also been a victim of the cuts and are now required to pay full tuition fees — up to \$5,000 a year

— unless they come from one of the 44 earmarked developing countries.

Against this background, it is surprising that the peculiarly Belgian system of "dedoublement", of maintaining separate university systems for each language, has survived intact. There is thus a Dutch-language university in Brussels (FUB) and a French one, each offering roughly the same courses. Similarly there is the old Catholic University of Louvain and the Flemish KUL at a modern campus at Louvain La Neuve. With a few exceptions on the humanities side, no courses are held jointly and teaching posts are awarded on a strictly linguistic basis. A Walloon physics professor would be unlikely to obtain a vacant post at a Flemish university.

The problem is further complicated by the religious requirements until recently in force in some Belgian universities. Filling a chair in, say, astrophysics, with a Dutch-speaking Catholic who happens also to be a reasonably distinguished astrophysicist is not always easy. The feasibility of reorganization within the university system is further complicated by the emergence of separate language-based regional governments, likely to resist moves threatening their "own" universities.

The result is that in post baby-boom Belgium, there are now some 15 universities for a population of 9.9 million. For the 4.4 million French speakers in Belgium, there are nine universities. The

Catholic University of Louvain — Dutch speakers need not apply.

dedoublement, says Andre Degroeve, president of the administrative council of ULB, is definitely not a cause of the financial problems of Belgian universities, which stem from a drop in the number of students and the need to reduce public spending.

The plans put forward by the affected universities consist chiefly of cutting back on salaries and staff, and the universities have been given *carte blanche* by the government to ignore existing employment contracts. Liège, for instance, will enforce the compulsory retirement of teaching staff at the age of 65 and of other employees at 60 from September 1983. The aim will be to reduce personnel by 600, some 25 per cent of the total, and to cut running costs by 12 per cent. Research assistantships will be particularly badly hit in all the affected universities. At ULB, three out of five jobs will disappear and Degroeve predicts that in ten or fifteen years there will be a serious shortage of experienced teaching staff.

Research work will also be badly affected and Belgian universities have not so far been especially successful in winning contracts from industry. "What makes the future look all the more dismal is that nobody knows how long the coalition government and its policies will last. Before the last elections there were six coalitions in four years", said Degroeve mournfully.

Jasper Becker



Polish Academy of Sciences

Foreign travel rights restricted

Scholars employed by the Polish Academy of Sciences who wish to travel abroad will from now on need the approval of "the appropriate party secretary". According to a recent circular from Dr Zdzisław Kaczmarek, academic secretary of the academy, this change is designed to prevent the issue of passports to "persons whose socio-political activities and attitudes" are "negative" or "openly aimed against the constitutional, and legal order obtaining in the Polish People's Republic".

The circular forms part of the instructions concerning the process of political "verification", now being implemented in two stages within the academy. The restriction on travel has ostensibly been imposed in order to safeguard the prestige of Polish science abroad and to improve the international image of Poland. It is, however, a marked reversal of recent academy policy on foreign travel. Two years ago, after the signing of the Gdansk accord which inaugurated Solidarity and a nationwide wave of liberalization, the academy demanded that it alone should be the arbiter of whether or not its members and employees should travel abroad. Such a clause, it was hoped, would be written into the promised new legislation on the Academy of Sciences; but martial law was imposed last December a few days before the session of the academy that would have debated this and other liberal clauses including proposals to change the role of the academic secretary. Indeed, Dr Kaczmarek's instructions on verification are an ironic reminder of what that session had hoped to achieve. For the academic secretary, whose post is of ministerial rank under the present structure of the academy, is responsible, in the first instance, not to the academy members or council but to the prime minister.

The new restrictions on foreign travel — and indeed the whole verification process — could affect members and employees of the academy deeply. In Poland, the Academy of Sciences is organized completely separately from the universities (unlike, for example, the Soviet Union, where academicians also hold university lectureships) and since the university purges of 1968, the academy has been a haven for scholars whose political views were considered too suspect for them to be entrusted with the teaching of young people. Indeed, the leading dissident group of the late 1970s, the "Committee for Social Self-Defence" (KOR) included among its 33 members two Academicians (Drs Edward Lipinski and Jan Kielanowski), while several others were employed in various academy institutes.

Objections to the new clampdown on travel could, however, in the current

draconian climate, well be taken as further proof of dissent. A recent attack in the party media alleged that cultural and scientific contacts between Poland and the West had been exploited by Western intelligence agencies. At the end of the 1970s, it is claimed, foreign scholarships were given to persons selected as future opposition leaders, while agents were sent to Poland disguised as "scholarship holders, lecturers, correspondents and experts".

Vera Rich

UK universities

Modest hiring prospect cheers

British universities seem this week to have been enormously cheered by the prospect of recruiting some hundreds of younger scientists and engineers to academic posts. The only fly in the ointment is that this prospect is made possible only by the £6 million removed by the Secretary of State for Science and Education, Sir Keith Joseph, from the estimated budget of the Social Science Research Council over the next three years. One research council official confessed earlier this week that the "new blood" money seemed a little like "blood money".

As yet, it seems not to have been decided how much of the £6 million will be available in the coming financial year, a decision that Sir Keith has left to the Social Science Research Council. Nor has it been decided how the funds thus liberated will be shared among the other four research councils, although the Science and Engineering Research Council, the largest and with a direct responsibility for university research, seems certain to get the largest share. The chances are that the funds available will be shared out among disciplines, and that departments will be allowed to compete for the posts available by telling hard-luck stories to the councils.

While the numerical effect of the £6 million that will be available over the next three years remains to be determined, the research councils are already congratulating themselves on having been able to counteract some of the damage done to university research by the financial pressures on university budgets.

Thus the Science and Engineering Research Council has supported at least 23 applications under its Special Replacement Scheme, under which it will pay the salary of a senior researcher for up to five years provided that the university will undertake to appoint a younger researcher to a permanent academic post immediately. Applications under this scheme appear, however, to have tailed off in the past

several months, apparently because of the counter-attractions of the government-backed early retirement scheme.

So far, there appear also to have been few takers for the scheme under which the council will pay the salaries of researchers left stranded by the reorganization of universities involving the closure of departments. Two such awards have so far been made, both of them to casualties of the decision of a year ago that the Department of Biological Sciences at the University of Bradford should be "phased out".

Dr Michael Lord, a biochemist concerned with cell organelles, has joined the chloroplast group at the University of Warwick (now renamed plant biochemistry). The council will pay Dr Lord's salary for the next ten years, while the four members of his group transferred from Bradford will continue to be supported from research council funds. Similarly, Professor M.J. Merrett has moved to the University of Wales at Swansea with the help of a research council subvention to the university. ●

Cancer research

French change

Jack Ralite, French minister of health and one of the few communist ministers in the Mitterrand government, is contemplating a major reorganization of cancer research and the care of cancer patients, to judge by his moves earlier this month.

First, he has dissolved the "high committee" on cancer, which was only a little over two years old, essentially because it had failed in its task of uniting the two major French cancer charities, the Ligue Nationale Française contre le Cancer and the Association pour le Développement de la Recherche sur la Cancer à Villejuif (ADRCV). These charities control budgets amounting to more than FF 122 million (£10 million), considerably more than government expenditure on cancer research. But because of personality conflicts, the two have never been able to develop complementary research policies. The "high committee", now abandoned, was the previous government's attempt to knock their heads together.

The Ligue and ADRCV may now have only a temporary respite, however. For the minister has also launched a great national enquiry on cancer (the Concertation Nationale sur le Cancer) which is designed to provide the basis for a new politics of cancer.

Through the enquiry, which is addressed to all concerned, the minister will no doubt learn how French cancer research and care is riven by factionalism and the philosophy of the *savant*. He may also be expecting to confirm some other more debatable ideas about research, which come through clearly in the questionnaire which is the basis of the Concertation Nationale.

The questionnaire proposes that of the three principal sectors of researchers in cancer — biologists, clinicians and epidemiologists — the biologists are “very preponderant”. Is it not important today, asks the questionnaire, to reflect on the evolution of this state of affairs “and on the consequences of this primacy given to reductionist approaches?” Biologists tend naturally to choose areas of research as close as possible to fundamental biology, says the questionnaire. “Is this not to the detriment of specific research on cancer?” And, the questionnaire asks, should there not be separate rules for making evaluations of cancer research, so it is not subject to the more fundamental criteria applied by the medical research council (INSERM) and the Centre National de la Recherche Scientifique?

Such questions have been asked before, notably in the United States, but they are still appropriate and they are expected to produce some vigorous replies in France, with sectional interests being defended in every direction. Once the fur is flying, and the political positions of every sector are evident to all, M. Ralite will make “coherent recommendations”. This is democracy.

Robert Walgate

Nuclear war games

No survivors?

Moscow radio last week broadcast a virulent attack on Dr Edward Teller, the American physicist, for having published in *Readers' Digest* an article which suggested that nuclear war would not obliterate all life on this planet.

The broadcast, scripted by the official news agency TASS, alleged that Dr Teller has a vested interest in the arms race, since its “reversal” would deprive him of his “cushy job” as a presidential aide. Teller’s article, TASS alleged, was intended to “dampen the truly immense public interest all over the world” in the Soviet peace proposals tabled at the the United Nations.

The Soviet proposal for a total test-ban treaty to create a more favourable atmosphere for arms limitation negotiations ran into considerable opposition from the United States, which maintains that the problems of verifying compliance with any such agreement should be solved before negotiations are started. In fact the Soviet Union put forward a package of verification proposals, including the international exchange of seismic data and the right of signatories to any such treaty to demand an on-site inspection of any suspected explosion in the territory of a fellow signatory. Since the Soviet proposals were backed up by carefully orchestrated panegyrics in the Soviet press the TASS writers may have genuinely assumed that the US rejection of these initiatives has a similar media back-up. Teller’s role as “father of the hydrogen bomb” is stressed and the activities of the

Pugwash movement ignored. His attempts to refute some of the more fantastic myths of what nuclear war would entail, his quoting of instances from Hiroshima and Nagasaki (“bridges were open to traffic a day after the blast, trains ran on the second day and streetcars were operating on the third”), and his suggestions on how to remove radioactive contamination are construed, by TASS, as “cynicism and misanthropy”.

In fact, Teller’s thesis that there would be a significant number of survivors of a war involving nuclear weapons and other “means of mass destruction” has tacitly been accepted by Soviet civil defence

Asbestos hazard

Bankruptcy protects the rich

Washington

The question of liability for the long-term latent occupational diseases of workers has been further muddled by the gigantic Manville Corporation, a mining and manufacturing company that is a pillar of US industry, which has filed for bankruptcy to protect itself from claims from asbestos workers, which the company says, could total as much as \$2,000 million.

The company’s action in August threw into turmoil the question of how US corporations deal with occupational health problems, sending other asbestos companies to their lawyers for advice and making claimants and their advocates suspect a fraud. It will be months before the full implications for the declining US asbestos industry, and for other industries with similar problems, are known.

Manville’s action was an extraordinary use of the US bankruptcy courts; companies normally file for bankruptcy when their debts have become intolerable in order to prevent outright liquidation. In those circumstances, the court takes charge of settling a corporation’s debts, freezing creditors’ claims and other lawsuits and dividends paid to shareholders, while a reorganization plan is being agreed.

But Manville is a relatively healthy corporation, although it reported a loss for the first half of 1982 of \$32 million. Its executives believe that the effect of filing for bankruptcy will be, ironically, to make the company still healthier.

Manville, which is the largest US producer of asbestos products used in building, insulation, brake lining, shipbuilding and underground pipes, has 16,500 claims outstanding against it. Its decision to file for bankruptcy was prompted by a study indicating that another 32,000 claims might yet be made, with an average settlement cost of \$40,000.

The asbestos industry in general faces a tidal wave of claims, because there are an estimated nine million people in the United States who have worked in an asbestos-related industry and who might claim damages on account of the risk of develop-

planners. Recent reports of civil defence exercises, including practice evacuation of schools and hospitals, speak of the need for further training in the use of “means of protection against weapons of mass destruction”. A broadcast on Vilnius radio last summer by a civil defence official went into considerable details of the various gamma and neutron radiation monitors available to establish contamination of the population and of livestock, machines and equipment — suggesting precisely the same swift return to quasi-normality after nuclear bombardment for which Teller was castigated by TASS.

Vera Rich

ing mesothelioma or asbestosis, maybe many years later.

The law so far has not ruled in favour of the asbestos industry. A lawsuit decided by the Supreme Court of the state of New Jersey last month ruled out as a possible defence a company’s ignorance of possible adverse health effects at the time its workers were exposed to asbestos. Awareness of the health hazards of asbestos has been relatively recent. Manville itself (then known as the Johns-Manville Corporation) lost a case in a California court and was forced to pay damages of \$28,000 to a victim over and above the workman’s compensation.

The Manville Corporation has also disputed for seven years with its insurance agents the point at which their liability should begin: at the time when the workers were exposed, when the disease appeared or from the time when medical hazards existed. As a result of the filing for bankruptcy, claimants will be able to obtain from the corporation data about what it knew and when. One court has already concluded that Manville concealed knowledge of asbestos hazards. Nevertheless, the corporation has spent \$24.5 million lawyers’ fees alone, and has so far paid out \$24 million in damages.

“CONGRATULATIONS! YOU ARE NOW FINANCIALLY AS WELL AS MORALLY BANKRUPT.”



By filing for bankruptcy, the corporation can put off paying existing creditors until the courts decide what should be done. In the meantime, no new claims can be filed. In reorganizing itself,

the Johns-Manville Corporation assigned to the new Manville Corporation only those assets associated with asbestos, thus shielding 74 per cent of its total assets from future litigation. It has even argued that claims against the Johns-Manville Corporation are not valid because that corporation no longer exists.

Other companies have been doing much the same. The Union Asbestos and Rubber Company became Unarco Industries Inc. and then UNR Industries. It has not made asbestos products since 1962, but claims are still pouring in. When it created UNR Industries Inc., it assigned to that company only its pre-1970 assets, which amounted to \$24 million. (Its assets today are \$200 million.) Then UNR Industries Inc. filed for bankruptcy in Chicago on the basis of the intolerable outstanding asbestos claim, so perhaps inspiring the Manville Corporation to do the same. **Deborah Shapley**

NRDC and NEB

Union delayed

The National Research Development Corporation is still coining money from the cephalosporin antibiotics, synthetic pyrethroid insecticides and half a dozen other inventions, but its intended marriage with the National Enterprise Board, announced in July 1981, has been delayed. Sir Freddie Wood, chairman of both organizations (which also now have identical boards of directors), said last week that he hopes to know more about the British government's intentions in a few months, but that legislation is unlikely to see the light of day before the end of the next year-long session of the British Parliament, beginning next week.

Meanwhile, the common opinion is that a formal merger of the two organizations would indeed bring the benefits advertised for it. The corporation, set up in 1949 to exploit inventions arising in the public sector (including universities) has been run principally by technical people, and during its chequered history has been accused of neglecting commercial opportunities. The board, by contrast, is stronger on accountancy and financial management.

The corporation as a single entity seems in good financial shape, to judge from the annual report published a week ago. It earned £26.2 million in the year to 31 March, chiefly from royalties on licensed inventions. Payments to inventors (and their institutions) amounted to £3.53 million, rather less than the sum set aside by the corporation for taxation before reaching its net profit of £5.02 million.

In the years immediately ahead, however, the corporation's income is likely to decline as patents on the cephalosporins run out. The annual report says that synthetic pyrethroids should help bridge the gap, but that these insecticides are unlikely to be as profitable. The cephalosporins stem from research at the University of Oxford, the pyrethroids from the Rothamsted Experimental Station.

During the year, the corporation says that 229 patents were assigned to it, a small increase on the previous year — and a surprisingly small number absolutely. The total of inventions about which the corporation was told amounted to 1,935, well over half from the public sector. Something like 350 inventions were earning money at the end of the year, while £23 million had been committed (or spent) on the development of patented inventions and £48 million to joint development ventures with industrial companies. ●

Astrophysics laboratory

Theoretical gain

Chicago

The Fermi National Accelerator Laboratory (Fermilab) has set up the first theoretical astrophysics group to reside at any of the world's accelerator laboratories. The group will start work in January, and is being supported by the US National Aeronautics and Space Administration to the tune of \$500,000 over three years. By January, Dr Leon M. Lederman, director of Fermilab, expects to have taken on one or two senior researchers and two or three junior workers.

Lederman says: "One of the most exciting parts of particle physics research now is the connection between inner and outer space. Particle physicists look down with their microscopes and astrophysicists look up with their telescopes and find they are looking at some of the same things".

Fermilab's studies are essential to model the evolution of the Universe immediately after the Big Bang, but Lederman says that the behaviour of the early Universe in turn constrains what the particle physicists can expect to find as accelerator energies increase.

The new group, with a huge accelerator readily available, will deal primarily with the connection between particle physics and the early Universe, in which high temperatures and energies were generated by the Big Bang. But current interests in particle physics research, the grand unified theories for example, deal with particle energies around 10^{14} GeV, ten orders of magnitude greater than the energies being dreamed of by today's accelerator builders.

Yet astrophysical theory can extrapolate as far back in time as 10^{-43} seconds after the Big Bang — when the thermal energy would have been 10^{19} GeV. The early Universe is thus an ersatz particle accelerator for extreme energies. The group will also investigate the implications of particle physics theory for astrophysics. Some grand unified theories suggest that the Universe passed through an "inflationary" phase of violent expansions that may explain why the Universe today is isotropic, homogeneous and dynamically "flat".

Lederman says that Fermilab has hitherto only "dabbled" in astrophysics, and he hopes that the new formal collaboration will enable more systematic and fruitful studies to be made. **Larry Arbeiter**

Celltech eyes Japanese market

Celltech, the British biotechnology company, has concluded a major five-year agreement for exclusive distribution of its alpha-interferon purification and assay products with the Sumitomo Corporation of Japan. The deal also covers Celltech's monoclonal antibody ABO blood grouping reagents, as well as anti-beta and gamma interferons.

Sumitomo, one of Japan's largest trading companies, has been selling Celltech products on an informal basis for several months; now Celltech hopes to reach a dominant position in the Japanese pharmaceutical and diagnostics market, valued at £8,500 million per year. Celltech is planning to develop monoclonal antibodies that can be incorporated into diagnostic kits and also hopes to gain some contract research from the present agreement; other cooperative ventures are not excluded.

The agreement follows closely that between Biogen of Switzerland and Fujisawa for joint development of

Biogen's human tissue plasminogen activator, which has yet to reach the stage of clinical trials. Hopes are high for this anticoagulant protein, which is more specific in its effects than other thrombolytic agents now in use, urokinase for example. It may come to represent a significant fraction of the large Japanese market for such agents. Fujisawa already has agreements with overseas companies for the production of human interferons.

These agreements follow a trend which has become apparent over the past few years in which Japanese firms arrange joint development and distribution deals with European and US companies; such arrangements have been reached between Hoffmann-La Roche, Genentech and Takeda, for example. While the Japanese appear actively to be developing their own expertise in techniques such as the production of monoclonal antibodies and recombinant DNA, they seem also to be encouraging the import of foreign products. **Tim Beardsley**

Ciba Foundation

In *Nature* for 30 September, p.383, there was a report of what Dr R.G. Edwards had said at a Galton Lecture and at a meeting organized by the British Medical Journalists' Association; this meeting was sponsored by Ciba-Geigy Pharmaceuticals and not, as reported, by the Ciba Foundation.

Social sciences in France

Few friends on the right

The unsympathetic manner in which Britain's Sir Keith Joseph regards the social sciences has precedents across the Channel, according to a long-awaited report on the social sciences in France. Figures collected by M. Maurice Godelier, a respected left-wing anthropologist and friend of the Mitterrand government, show that Giscard d'Estaing, the previous President of France, slashed social science spending by more than half between 1976 and 1981, and the humanities by nearly a quarter. Now is the time for reconstruction, and the government seems likely to follow M. Godelier's recommendations, although science and industry minister Jean-Pierre Chevènement said on introducing the report that full implementation would be too costly.

According to Godelier, there is much work to do. Giscard caused such a collapse in the human sciences "that the general public would hardly believe it". Libraries and documentation centres were particularly badly hit. The result was that subjects not protected by this or that "mandarin" were smothered. In the scramble for funds, academic standards went by the board. Economics fared best, but only mainstream neoclassical economics. Sociology, tainted with being "leftist", was worse hit.

Godelier is thus starting almost from scratch. He recommends recreating infrastructures destroyed by Giscard (contract research supported by government departments had almost disappeared). He wants buildings repaired, libraries re-equipped, more funds for field studies and for publi-

cations, improved international contacts and more staff.

Nobody is likely to disagree with any of that — except for the size of the bill. But eyebrows have been raised in some quarters by his statement that the "new dynamic" should be founded not only on an increase in funding but also on a different cutting of the cake, in response to "a new evaluation of needs". These, it seems, would in part be turned towards the realities of modern French life, involving the study, for example, of both business and administration — the elite.

Moreover, Godelier has recommended the setting up of many more interdisciplinary studies, and cross-border funding committees — which might threaten the very "mandarins" who survived Giscard. Godelier is thus seen as a personal and political threat in some quarters; and when this is combined with his forthright manner, he finds himself facing considerable opposition.

Thus when Chevènement attempted to make Godelier director of social sciences at the Centre National de la Recherche Scientifique (CNRS) last year, key figures in CNRS resigned over this alleged interference with "academic freedom". But others have pointed out that the man Godelier was to replace (and who did indeed resign, leaving a vacancy still unfilled) was himself a political appointee, a friend of Giscard's prime minister, Raymond Barre. Chevènement's logic was not faulty.

The political problem now is to fill the

hot seat at CNRS, from which the Godelier reforms will be implemented, without creating too much uproar. Some kind of compromise, which might perhaps involve Godelier getting half the job but not the full responsibility, seems to be in the offing.

Robert Walgate

Education in Afghanistan

Russian aid

Afghanistan has embarked on a major programme of expansion of higher and technical education, Mr Sarwar Mangal, the Minister of Higher and Vocational Education, announced recently. Mr Mangal, who became minister at the end of September, was taking part in a series of special broadcasts on Kabul radio in which ministers explain the work of their ministries.

The expansion programme will concentrate on those areas most important to the national economy — in particular engineering and agricultural sciences. Three new agricultural technical institutes will be founded at Mazar-e Sharif, Jalalabad and Baglan, and the agricultural secondary school at Helmand will be upgraded to an agricultural technical college. Major building works are under way at Kabul University (a chemistry institute, classrooms block and accommodation for lecturers and students), Nangarhar University (faculty of education and agriculture school) and Balkh (agricultural school). A new accommodation block for Kabul Polytechnic Institute will be started in the near future. Mr Mangal admitted, however, that there were "deficiencies, shortcomings and improper management" in the building programme.

Higher education in Afghanistan has mushroomed in the past few years. Enrolment in higher education (including the Kabul Polytechnic Institute and the Mangarhar and Kabul medical schools) in the Afghan year 1361 (March 1982–March 1983) was 155, a 250 per cent increase compared with 1355 (1976–77). Entrance examinations were recently standardized, although ex-soldiers are admitted to university without taking the competitive examination. Recruitment to technical-vocational education seems to be going more slowly; Mangal spoke of recent improvements in the vocational sphere, including the launching of poster campaigns and sports contests to attract young people to vocational courses. The first "professional-technical" school, which can train 100 students in the maintenance of vehicles and industrial machinery and which was constructed under a grant-in-aid from the Soviet Union, was opened recently, and a similar television and refrigerator maintenance school, equipped by the Soviet Union, should be ready at the end of this (Afghan) year or the beginning of the next.

Spending low on France's books

A library, to a researcher in the humanities or social sciences, writes Hervé Le Bras, director of research at the Institut National des Etudes Démographiques, is like a telescope to an astronomer. But, claims Le Bras in an appendix to the Godelier report (see adjacent story), French libraries (though never good) are now "catastrophic".

The present state of affairs is illustrated by a comparison between the British Library (based in part at the British Museum) and the French equivalent, the

Bibliothèque Nationale. The table shows, in particular, a lack of French funds, posts and lending.

Some other figures: the Bibliothèque Nationale buys only 3.6 per cent of European titles in languages other than French, and only 2 per cent of North American titles. "The shadow which covers our own history [through lack of texts and documents] also extends to the rest of the world" says Le Bras. French social sciences "must not become provincial" he says.

	British Library (1979–80)	Bibliothèque Nationale (1981–82)
Total budget (FF)*	381,205,000	61,740,766
Conservation budget	24,200,000	11,611,508
Acquisitions of titles	2,100,000	750,000
Accredited readers	18,800	15,600
Staff	1,125	802
Requests for borrowing	2,375,000	24,700
Lending for borrowing	2,375,000	24,700
Lending budget (FF)	22,792,000	250,000
Lending staff	730	54

*(FF 12 = £1)

Higher and professional education is critically dependent on Soviet aid. As well as supplying funds for the professional schools, the Soviet "Progress Publishing Company" is to produce more than 200 university textbooks with a total print-run of 300,000 copies as grant in aid. (These will mostly be translations of existing Soviet texts). Moreover, although postgraduate courses up to master's degree

are now available at the Kabul Polytechnic Institute and will shortly be introduced at Kabul University, most postgraduate training can still take place only in the Soviet Union. This year, some 1,500 graduate students began courses in Moscow, which, compared with the 4,155 freshmen enrolled this year, constitutes a sizeable proportion of the student body.

Vera Rich

FDA on overseas data

US drug market to open up?

Washington

New drugs may be approved for sale in the United States solely on the basis of foreign data under a proposed reorganization of the Food and Drug Administration (FDA) new drug application process. One effect may be to facilitate the entry of drugs developed abroad into the US market.

The proposed changes, published last week, have already provoked some strong reactions. Under current FDA rules, at least one of the clinical studies of a new drug must be conducted within the United States. Critics claim that FDA's proposal, increasing the admissibility of foreign data, flies in the face of evidence that foreign studies are difficult to verify and that standards for protection of human subjects are generally lower outside the United States.

In a discussion that accompanies the new proposal, FDA concedes that there are problems in accepting foreign studies. For one thing, genetic differences between foreign and US populations may render foreign results inapplicable; similarly, differences in medical practice and even in terminology (the definition of "depression", for example, is shaded by cultural differences from country to country) may also be significant. FDA admits, too, that the competence of foreign researchers is more difficult to judge than that of US scientists. FDA's proposed solution is to reject applications based solely on foreign studies when "the calibre of the key clinical investigators and facilities is unknown" or when there is "reason to believe" that genetic, medical or cultural differences affect the applicability of the results to the United States.

But a potentially more serious obstacle to the acceptability of foreign data is the difficulty of auditing foreign clinical trials. A House of Representatives subcommittee found last August (see *Nature* 12 August, p.598) that FDA investigators were unable to gain access to the medical records of a clinical trial in a Canadian hospital because of local confidentiality laws; and an audit of one Mexican study found patients' records destroyed.

FDA says it will reject applications if a "for-cause" inspection is considered necessary and then cannot be carried out

because of such obstacles. But Dan Sigelman, who is on the staff of the House subcommittee that investigated FDA, says that this is a "Catch-22". FDA routinely conducts spot-check audits of domestic studies and these, according to Sigelman, are what normally turn up the cause for further "for-cause" inspections. "On what ground are you going to determine the need for an audit without an audit?" he asks.

Sigelman also questions FDA's assertion that US drug companies will continue to favour US studies so that they can acquaint US physicians with the new drug before it is marketed. "They can sit and argue all they want that drug companies won't go abroad, but if the trade-off is getting a drug on the market more quickly, you're just opening up the floodgates to foreign data", he says.

The change on foreign data had been pressed by the US drug companies' trade group, the Pharmaceutical Manufacturers' Association (PMA). PMA also got its way on another controversial point: FDA is proposing to drop the current requirement that case-report forms from clinical trials be submitted with new drug applications. In place of these forms, which are made out by the clinical investigator on each patient, a tabulation of the raw data would be submitted. Under this proposed change, FDA could still request the case-report forms, but only when a "legitimate need for them exists in order to conduct an adequate review of the application".

The only noticeable tightening of the rules in the FDA proposal concerns the reporting of adverse findings about a drug by its manufacturer. The current rules are vague on how and even whether adverse findings are to be reported to FDA once an application is on file for approval — as the case of Eli Lilly and Co.'s reporting of deaths among Oraflex (benoxapofen) users demonstrated earlier this year.

FDA's proposal requires drug companies to file a safety report every four months when it has an application on file. FDA is also proposing to tighten its requirements on reporting of adverse effects of drugs already on the market: fatal and life-threatening effects would be reported to FDA within 15 days in "alert reports", other "adverse experiences"

within 30 days.

Critics are calling these changes window dressing, however. Dr Sidney Wolfe of the Ralph Nader Health Research Group says, "my response to this whole stunt is that the more important issue is enforcing existing regulations". Wolfe cites FDA's continued failure to bring criminal charges against Lilly, as recommended by a former FDA investigator, for withholding adverse effect data on Oraflex and three other drugs.

FDA is accepting public comments on its proposed changes until 20 December. After digesting these — and possibly making some alterations — the agency will publish a final rule, probably in early spring. As the proposals stand now, though, it is clear that the big winners are the drug companies, which will be able to file less paper, receive quicker responses and have an easier time introducing drugs already marketed abroad into the United States.

Stephen Budiansky

Halley upstaged?

Washington

The US National Aeronautics and Space Administration (NASA) has set one of its spacecraft on a complex manoeuvring course that will enable it to intercept and study the Giacobini-Zinner comet on 11 September 1985 — six months before Soviet, European and Japanese spacecraft are due to meet Halley's comet.

The Giacobini-Zinner comet, which approaches the Sun every 13 years, will not be visible from the Earth, but the better known Halley's comet has already been detected with the 5-metre Hale telescope on Palomar Mountain.

NASA's plan is to move the International Sun-Earth Explorer (ISEE 3), which has been in a permanent orbit between the Earth and the Sun since 1978, measuring electric and magnetic fields. US scientists have been upset by the Reagan Administration's cancellation of the \$250 million plan for a US spacecraft to Halley's comet. The NASA decision to use ISEE 3 to intercept another comet first may console them, because the United States will thereby be the first to provide valuable data that others can use in analysis of Halley's comet.

ISEE 3 has already been moved to the side of the Earth away from the Sun and on 6 February next year it will be directed on a course that will take it past the Moon. It will then be brought close to the Moon to use its gravity to give the spacecraft a push towards the comet. After the Giacobini-Zinner probe, ISEE 3 may be used to measure the solar wind extending from the Sun towards Halley's comet at the time when the other probes reach that comet early next year.

Deborah Shapley

Nuclear fuel supply

World enrichment over-capacity

Brussels

The Euratom Supply Agency, which has the sole right to negotiate nuclear fuel supply contracts for Euratom's 10 member states, points out in its annual report for 1981 that EEC is in danger of creating an overcapacity of nuclear fuel enrichment services. Last year EEC registered a positive trade balance of 2,200 tonnes SWU (separative work units) compared with its own requirements of 4,500 tonnes. By 1984 the European Community could have a capacity of 11,800 tonnes and an internal demand of 7,400 tonnes.

The over-capacity is likely to become more of an embarrassment, since Japan and Brazil, which until now have been in the market as buyers, have decided to build their own enrichment plants. What is worse is that Australia, which has signed a long-term supply contract with EEC, plans to upgrade locally the uranium it mines, and detailed studies are already under way for the construction of a plant for isotopic separation. This may create political problems as some of Euratom's member states would prefer to keep separate the purchase of natural uranium and enriched uranium for security reasons.

By 1990, according to the Euratom

Supply Agency, world enrichment capacity is likely to be around 49,650 tonnes SWU compared with the 1980 figure of 36,150 tonnes. The United States still dominates the industry with a 27,300 tonne capacity but from 1989 onwards France will be capable of enriching 13,300 tonnes.

The prospect of competition from France may have a significant impact on the costs of enrichment. The price of the US Department of Energy's enrichment services is steadily increasing, reports the Euratom agency. The cost per SWU rose by 27.4 per cent last year (from \$111.75 to \$141.14) and the Department of Energy announced a further price rise for 1982. These price rises have been further exacerbated by the decline in the value of the European Currency Unit (ECU) against the dollar so that enrichment costs at the "front end" of the fuel cycle now exceed 40 per cent of the total cost of operations while the natural uranium component has for the first time fallen to below 40 per cent of the cost. Uranium users have therefore been hanging onto their stocks of natural uranium rather than enriching it prematurely.

The price of plutonium has also fallen, from \$10 to \$4 per gramme of Pu fissile

Uranium enrichment in Euratom nations (tonnes SWU)

Production	Requirement	Balance
1978 —	1,400	-1,400
1979 2,600	3,200	- 600
1980 6,000	3,900	+ 2,100
1981 6,700	4,500	+ 2,200
<i>Capacity</i>		
1984 11,800	7,000	+ 4,800
1985 11,800	7,400	+ 4,400

material, and the Euratom agency predicts that in future the supply of plutonium could considerably exceed that required by fast reactor programmes. Similarly, natural uranium is also likely to remain a buyer's market with an excess of supply over demand. Prices have remained stable, with world stocks of natural and enriched uranium now equivalent to more than three years' consumption. Deliveries of natural uranium to Euratom countries amounted to 13,000 tonnes in 1981 and this is likely to drop to 10,750 tonnes by 1983.

Jasper Becker

French nuclear power

EDF's setbacks

Electricité de France (EDF) is having increasing trouble with some tiny but important parts of its pressurized water nuclear reactors, which were due to provide half of France's electricity by next year. In fact the problem is so worrying that EDF says it will have to close down 20 stations for repair between 1983 and 1984, with a total loss of some 22 operational reactor-months.

The pieces are clips which hold up guide tubes for the control rods — rods which must be driven or dropped into the reactor core to close down the nuclear reactor in case of accident. In January, pieces of one of these clips were found lodged in an emergency cooling pipe in one of the reactors at Gravelines, and since then other broken clips have been detected at Fessenheim and in two reactors at Bugey. The loss of the clips is potentially dangerous — they could block the proper entry of the control rods, or, by circulating in the cooling system they could damage pumps and valves.

Moreover, it seems that the broken clips had no individual faults: the error appears to have come in the choice and treatment of the clip material (which is exposed to high temperatures and radiation levels in the reactor). That is why the clips must be replaced in all 20 reactors where the material was used, at a cost (mostly in lost electricity production) of at least FF 1,000 million.

The incident illustrates one of the dangers of the massive French nuclear programme, which reduced costs by using long series production of a few single designs. If the design is wrong, all fail at once. "This just shows how the French nuclear industry has all its problems yet to come" said one commentator this week. **Robert Walgate**

Projected world enrichment capacity (tonnes SWU)

	1981	1982	1986	1987	1988	1989	1990
Eurodif (France)	10,000	10,800	10,800	10,800	10,800	10,800	10,800
Urenco	550	750	2,100	2,300	2,700	3,000	3,300
DoE diffusion (US)	26,900	27,100	27,300	27,300	27,300	27,300	27,300
DoE centrifugation (US)	—	—	—	—	1,100	2,200	2,200
Technabexport (USSR)	3,000	3,000	3,000	3,000	3,000	3,000	3,000
PNC (Japan)	—	—	100	100	250	250	250
UCOR (South Africa)	—	—	—	—	100	100	100
Nuclei (Brazil)	—	—	100	200	200	200	200
Coredif (France)	—	—	—	—	—	2,500	2,500
Total	40,450	41,650	43,400	43,700	45,450	49,350	49,650

India embroiled in nuclear politics

Bangalore

After more than three months, negotiations between India and France over the supply of enriched uranium for India's troubled nuclear power station at Tarapur have reached stalemate. France is insisting on operating the safeguards laid down by the International Atomic Energy Agency (IAEA), aimed at limiting the spread of nuclear weapons. To the Indian negotiators, however, such restrictions imposed by Western countries on the supply of nuclear fuel "devalue India's national prestige" by asking for "humiliating terms and conditions".

The negotiations, described by India's Foreign Minister Mr Narashima Rao as "a tortuous process", did see some compromises made by France on the matter of reprocessing spent fuel in India. But these proved insufficient to break the

deadlock, with India being particularly unwilling to comply with France's alleged insistence on a written undertaking that India will not carry out any further testing of nuclear bombs.

The Indian government is also in disagreement with the Soviet Union, having declined a Soviet offer to build a 1,000-MW nuclear plant in India, despite its problems in obtaining enriched uranium for its Tarapur plant. The Soviet offer, made previously when Mr Morarji Desai was Prime Minister, was renewed during Mrs Indira Gandhi's recent visit to Moscow.

Like France, the Soviet Union is insisting on the IAEA safeguards. In the face of restrictions from both East and West, members of India's parliament have called for a speeding up of plans to develop MOX, an indigenous fuel supply.

B. Radhakrishna Rao

CORRESPONDENCE

Ball lightning in the laboratory

SIR — Sir Brian Pippard's account of ball lightning observed outside the Cavendish Laboratory¹ is the latest of many reported sightings of this phenomenon. Unfortunately, Nature is not an obliging wench² and men generally experience her performance no more than once in a lifetime. Thus descriptions garnered from single momentary unexpected glimpses tend to be inexact and unscientific; if the fireball be the angular size of the Moon¹, what is its diameter? I would like therefore to draw attention to the possibility that some photographs we published³ some years ago were in fact of a laboratory version of ball lightning.

The work was a consequence of some peripheral observations we made while developing a streamer chamber⁴ for cosmic ray studies. Originally we discovered a ringing phenomenon superimposed on the Lichtenberg figures that appeared on the insulated bottom of the chamber and centred on the particle tracks⁵. This was with a chamber filled with the noble gas mixture of 70 per cent neon and 30 per cent helium. These dark rings could manifest themselves as dark spaces in a columnar discharge⁶ and the advance of the discharge could be studied by arresting the field and photographing the various stages⁷. It was when we extended the work to air³ that a new effect emerged which may now shed light on the nature of ball lightning.

The photograph shows the electrical discharge resulting from the application of a

avalanches but which here can only move upwards and so some photoionization mechanism must be assumed. The downward velocity of the luminous front is of the order of 5×10^5 m s⁻¹.

The curiosity here is the well-defined luminous cone seen in the photograph (region 3), for which we have seen no counterpart in noble gases. The photomultiplier trace⁸ shows that light is emitted synchronically with the oscillations of the applied voltage — even though the adjacent regions^{2,3} emit essentially only once when the discharge front traverses. The cone is seen to be a stationary, stable, localized fireball maintained by the electric field. If the field oscillations were allowed appropriately to continue for seconds, one might suppose the emission from the plasma cone would similarly continue.

The similarity between the suspended cone of plasma and ball lightning prompts the hypothesis that ball lightning itself is maintained by an oscillatory atmospheric electric field. The reason for the localized containment of the plasma is an unanswered question, but one which is now open to experimental study in the laboratory. Ball lightning moves gently and one must presume that this is in response to the site of the advantageous field conditions itself moving. In that discharge phenomena tend to be a function of the ratio E/p , the higher atmospheric pressure would imply higher electric fields.

P.C. RICE-EVANS

Department of Physics,
Bedford College,
University of London,
London NW1, UK

1. Pippard, A.B. *Nature* **298**, 702 (1982).
2. Tonks, L. *Nature* **187**, 1013 (1960).
3. Rice-Evans, P.C. & Franco, I.J. *J. Phys* **D13**, 1079 (1980).
4. Rice-Evans, P. *Spark Streamer, Proportional and Drift Chambers* (Richelieu, London, 1974).
5. Rice-Evans, P. & Hassairi, I.J. *Phys. Lett.* **38A**, 196 (1972).
6. Rice-Evans, P. & Franco, I.J. *Phys. Lett.* **63A**, 291 (1977).
7. Rice-Evans, P. & Franco, I.J. *Phys. Lett.* **70A**, 20 (1979).
8. Golde, R.H. *Lightning* Vol.1 (Academic, London, 1977).

ICSU matters

SIR — In view of the factual errors in "Doing one's thing" *Nature* 16 September, p.194, I should be grateful if you would publish the following:

(1) Discussions have been taking place with representatives of the China Association for Science and Technology and the Academy located in Taipei, China, for the past ten years to prepare the terms of an agreement that would be acceptable to all three parties and would make it possible for both organizations to be members of ICSU with the right to vote. These were successfully completed on 15 September, as Vera Rich indicates in your issue of 23 September.

(2) The communication from the Government of Australia to the Australian

Academy is quite specific and does not give "terms of reference wide enough, etc.". It did not add the proviso that "in future it should have a hand in planning conferences of this kind". To which kind of conferences do you refer?

(3) Sir John Kendrew was elected as first *Vice-President* for two years. He will become *President* in 1984 for four years.

(4) If ICSU is "short of money" it is because of a general shortage of money in the academic world, which provides most of the funds for ICSU, not because "it is trying to accomplish goals for which it was not created".

Further, in relation to "Academies not international" (23 September, p.288), the argument from which the predecessor of ICSU, the International Research Council (IRC) and ICSU spring is: the coordination of international efforts in the different branches of science and its applications.

Today ICSU's ultimate objectives include "to encourage international scientific activities for the benefit of mankind and so promote the cause of peace and international security throughout the world", and in pursuing these objectives ICSU observes a basic policy of non-discrimination — it affirms "the rights of scientists throughout the world to adhere to or to associate with international scientific activity without regard to race, religion, political philosophy, ethnic origin, citizenship, language or sex". The withdrawal of Taiwan was not an acceptable solution.

ICSU does not have and does not need a UN sponsor. As an organization expands, which ICSU has done in the past three decades (10 new International Scientific Unions, 26 new National Members, 17 new Scientific Associates, 4 new National Associates Biological Programme, the Upper Mantle Programme, the world Scientific and Technical Information Study jointly with Unesco, the Global Atmospheric Research Programme jointly with WMO, etc), paperwork expands . . . as has the number of printed pages about scientific discoveries in the past three decades.

"The state of affairs" about which your article says ICSU should be worrying seems to be pure politics: an area ICSU tries to avoid — "tries" is put purposively because it is not possible for any organization, not even a non-governmental organization, to escape from the effects of politics and governments' policies.

F.W.G. BAKER

(Executive Secretary)

International Council of Scientific Unions,
Paris, France

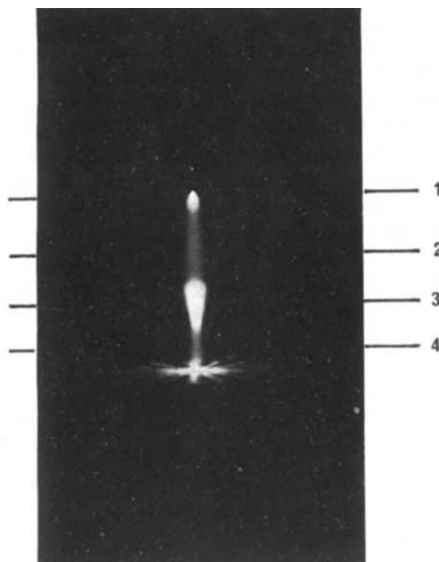
(1) Agreed; we said so twice.

(2) The Australian government has asked the Australian National Academy of Science to consult on future international conferences.

(3) We are grateful for this correction.

(4) A matter of opinion, linked with the question raised in the article "Science not international" of whether ICSU would be more effective in its laudable fight against discrimination if it were organized differently.

EDITOR NATURE



Photograph of the discharge corresponding to the positive high voltage pulse applied to the wire electrode (invisible here).

modulated positive potential to the wire with the lower plate earthed. The initial rise of the impulse causes field emission at the wire tip, and the plasma develops downwards. The precise nature of this advance is not understood; the main agents of discharges are electrons that participate in Townsend

NEWS AND VIEWS

Calculating the strength of wood

A group of mechanical engineers has provided a rational framework for understanding what is known of the strength of all kinds of wood. The methods should be more widely applied.

That wood is in some sense strong has been known since antiquity. Surprisingly, the reasons why have only in the past few weeks been made clear.

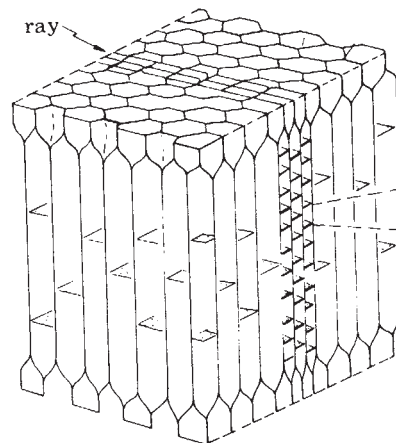
The construction industry probably includes few avid readers of the *Proceedings of the Royal Society*, either series A (physics) or B (biology), which is why something must be done to draw attention to the remarkable circumstances that this most august journal has now published what seems to be the first convincing explanation of the mechanical strength of wood. For what it is worth, the analysis (by K.E. Easterling, R. Harrysson, L.J. Gibson and M.F. Ashby, the first two from the University of Luleå in Sweden and the others from the Department of Engineering at the University of Cambridge) turns up in series A of *Proc. R. Soc.* (383, 31-41, 1982). The paper, which owes something to an earlier study of the structure of cork (L.J. Gibson, K.E. Easterling and M.F. Ashby, *Proc. R. Soc.* 377, 99-117; 1981) and to some newly reported investigations of the cellular structure of balsa wood, is a splendid illustration of what mechanical engineers can accomplish (if they are permitted) for biologists.

Living trees are of course aggregates of cells, most of them filled with cytoplasm, some of them seasonably capable of photosynthesis. The wood derived from trees that is used in the construction industry is, however, best thought of in quite different images, as a three-dimensional structure made from the walls of empty cells, themselves made largely from cellulose with a measured and nearly constant density of about $1.5 \times 10^3 \text{ kg m}^{-3}$, and with tensile strengths and other mechanical properties that can in principle be measured (but which in practice appear to depend on the orientation of the cellulose microfibrils within the cell walls). So, ever since the shapes of wood cells were first described by microscopists, it should have been a relatively simple task for mechanical engineers to figure out just how strong a particular wood should be. Alternatively, just as aeronautical engineers are forever trying to puzzle out what kind of honeycomb structure will provide the most resistance to some specified force (bending, twisting, crushing . . .) for the least use of material, so people should have been on the look-out for wood-cell structures promising desirable mechanical properties.

It will be interesting to see whether an attempt (not yet carried out) to predict the ideally strong wood would point to the cellular structure of a balsa wood now described — a three-dimensional network of cells with hexagonal cross-section with pointed ends and with an aspect ratio (of length to breadth) of about sixteen. In practice, the circumstances (see figure) are more complicated. In balsa wood, bundles of hexagonal needles oriented along the axis of the tree are lumped together around each annual growth ring, but are also separated from each other in blocks by layers of differently shaped cells running out in sheets from the centre of the tree (and which, when seen in the cross-section of a tree, not merely a balsa tree, are called rays). One important complication is that individual cells may have stiffening cellulose membranes perpendicular to the axis of the cell (and of the tree). What Easterling *et al.* have done is to calculate the mechanical properties of such structures.

The results of the analysis are surprisingly straightforward, but consistent with what has been found empirically about the strength of wood. Take, for example, the calculation of the Young's modulus of wood compressed tangentially. The only

way in which the wood cells can be deformed is by the bending of the angles by which the walls are joined, which can be related to the Young's modulus (in that direction) of the materials of which the walls themselves are made. Indeed, the tangential Young's modulus turns out to be the Young's modulus of the cell-wall material multiplied by the factor $4/\sqrt{3}(t/l)^3$ where t is the thickness of the hexagonal cell wall and l the length of the hexagonal side. Simple geometrical arguments lead to the conclusion that the tangential Young's modulus of balsa wood should be proportional to the cube of the physical density of the



Schematic representation of bundles of columnar hexagonal cells within a balsa growth-ring (from Easterling et al.) Cell height = 0.6 mm (approx), diameter = 0.035 mm.

wood. The same line of argument suggests that the radial modulus, corresponding to compression along a radius from the centre of the tree, is twice as large, but that the axial modulus should be proportional to the first power of the density of the bulk wood, not to its cube.

The remarkable feature of the comparison between prediction and observation now reported is not that the numerical values (calculated and observed) differ by factors which are sometimes as large as four or five but that they are in the same ball-park. Curiously enough, the mechanical engineers seem to be better able to predict the crushing strength of balsa wood, perhaps because the dominant factors in their calculations are estimates of the collapse strength of the exceptional pieces of the cell structure, the hexagonally prismatic pointed ends of balsa wood cells, for example. The outstanding achievement of the argument, however, about which more could with advantage have been said, is that the variation of the mechanical properties of balsa wood with density calculated from first principles provides scaling laws that appear to accommodate not merely this exceptional material but other kinds of wood as well.

That the scaling laws for the strength of wood should have a rational basis is comforting but unsurprising. What else would one expect? The interest of this study, however, is that it shows that in the investigation of biological materials (of which wood is one) even the macroscopic methods of mechanical engineering can be applied successfully to estimate the mechanical properties of microfibrillar structures, bundles of cellulose molecules and the like. Bone is another such material to which attention has now turned. No doubt there will be many others. □

The acetylcholine receptor cloned east and west

from Charles F. Stevens

A NEW era in neurobiology has begun with the successful cloning and sequencing of DNA that codes for subunits of the acetylcholine receptor. The α subunit sequence reported in this issue of *Nature* (p. 793) is another first for Shosaku Numa's group at Kyoto University and the sequence of the γ subunit, the successful cloning of which was announced recently², is expected to appear soon from the group at the Salk Institute, California³. The existence of cloned DNA coding for channel proteins will allow the power of molecular genetic approaches to be applied to the investigation of brain organization and function — it should be possible to categorize and understand the relationships between the numerous channels found in nerve cells and to see how they appear during development as well as to understand the molecular mechanisms of channel formation.

The acetylcholine receptor is one of a rather large family of integral membrane proteins that are responsible for the electrical activity of the nervous system. The proteins, known as channels, regulate the flow of ions across the membranes of excitable cells and can be divided into two broad categories — ligand-gated channels that are opened by the binding, with enzyme-like specificity, of a particular ligand and function in communication between cells, and voltage-gated channels² that are opened by voltage differences across cell membranes and are responsible for nerve impulses and transmission of information along axons. The acetylcholine receptor is a ligand-gated channel responsible for a link in the communication between nerve cells and skeletal muscle in vertebrates.

The acetylcholine receptor is the best understood channel protein, and has been extensively studied from physiological, biochemical and cell-biological viewpoints. It is a pentameric protein with a molecular weight of approximately 250,000 and is composed of four distinct polypeptide chains, with different molecular weights, designated α , β , γ and δ ; the α subunit is responsible for binding acetylcholine, the ligand that regulates ion flow through the channel, and two α subunits are present per channel. The two α subunits are very similar, but may not be identical.

In this age of cloning, what is notable about bugs growing the DNA coding for yet another protein? Three basic things can be done with the cloned DNA: it can be

used as a means for identifying, in hybridization experiments, other RNA or DNA molecules that contain the same, or similar, information; it can be sequenced; and it can be used to instruct cells to make the protein encoded by the DNA, possibly after modifying the DNA by mutations. All three paths are now available for the acetylcholine receptor.

The several dozen different channel species present in nerve cells are currently categorized according to the neurotransmitter ligand that regulates the ionic flow through them or, for the voltage-gated channels, according to the type of ion which passes most readily through. Thus, we speak of acetylcholine-activated channels, serotonin-activated channels, or sodium channels. Everyone believes that some order exists in the various channel types, but we have no way of categorizing channels now. For example, all ligand-gated channels may be related in some sense, or some subclasses, perhaps those activated by the biogenic amines, may be related. Furthermore, all channels that regulate the passage of a particular ion species (calcium or sodium, for example) may share certain structural characteristics. Presumably, then, channels will form one or several large families of molecules.

The key to understanding the relations within and between these families will be analysis of the genes encoding them. These genes, in turn, can only be studied when they can be identified and this is done with hybridization experiments using cloned DNA. Cloning the DNA for the acetylcholine receptor channel, then, is the first step in beginning to unravel the interrelationships between those protein molecules responsible for the electrical activity of the nervous system.

Understanding how the nervous system develops is one of the main challenges for modern neurobiology. Having DNA for receptor molecules also provides a tool for studying such important developmental questions as when receptors begin to be made and which receptors are made at which times during development. In developing muscle cells, for example, a species of acetylcholine receptor is inserted in membranes that is somewhat different from the one found in mature muscles. How and why is this receptor different, and what regulates its expression? Questions like these can be approached in hybridization experiments once receptor DNA is available.

Another group of questions of interest to modern neurobiology relates to the molecular mechanisms of channel function, and answering these questions

depends on sequencing channel DNA. What structural similarities are there between the various channels that permit sodium ions to pass through and excludes potassium ions? The structure of DNA coding for channels can give us the order of amino acids in the channel protein, and this in turn is the first step in relating structure and function.

A third experimental approach involves the manufacture of channel structures that have been modified by specific mutations of the DNA. By techniques similar to those used by McNight and Kingsbury⁴ one could, in principle, systematically alter small protein regions, and study the functional properties of the channel expressed in cultured cells or oocytes.

The Salk Institute group that has cloned the DNA coding for the γ subunit and the Kyoto University group used similar cloning techniques, and both present a sequence that has 5' and 3' non-coding regions and a 5' signal sequence characteristic of membrane proteins.

The structures of the two proteins encoded by the two DNAs are interesting. Protein sequencing data from Hood's laboratory⁵ had already shown that the first 50 or so amino acids in the four different acetylcholine receptor subunits showed strong homologies. The sequences for the α and γ subunits show that homology extends for about the first 150 amino acids and then is no longer evident. Both subunits share a probable site for glycosylation about 140 amino acids from the amino-terminal end. Despite the fact that the bulk of the protein has no apparent direct amino acid homologies, the overall shape predicted for the two polypeptides is remarkably similar: the amino-terminal end, presumed to be on the extracellular side of the membrane, is hydrophilic and is followed — for both subunits — by three hydrophilic sequences of about the same length which presumably loop through the membrane. A hydrophobic region, again possibly membranous, is found just before the carboxy terminus.

Although the overall procedures used by the two laboratories are similar, some differences are instructive. The Salk group used the approach of enriching their mRNA for the desired message, and then screened the cDNA clones by an *in vitro* translation technique. The Kyoto group screened a much larger number of clones with oligonucleotides synthesized from the known amino acid sequence at the amino-terminal end of the protein. This latter technique has the advantage of permitting many more clones to be screened, but the

1. Noda *et al.* *Nature* 299, 793 (1982).

2. Ballivet, Patrick, Lee & Heinemann *Proc. natn. Acad. Sci. U.S.A.* 79, 4466 (1982).

3. Claudio, Ballivet, Patrick & Heinemann *Proc. natn. Acad. Sci. U.S.A.* (in the press).

4. McNight & Kingsbury *Science* 217, 316 (1982).

5. Raftery, Huonkapillier, Strader & Hood *Science* 208, 1454 (1980).

6. Helfman, Feramisco, Fiddes, Thomas & Hughes *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Charles F. Stevens is in the Department of Physiology, Yale University, New Haven, Connecticut 06510.

disadvantage of being restricted only to those proteins where a probe sequence is available. It is probable that the use of expression vectors (see for example, ref.6) will offer a combination of the advantages of the two methods.

Now that the first step has been taken, we can anticipate rapid progress in

contributions by molecular genetics to understanding the organization and functioning of the brain. Quite probably, the first step will come through hybridization experiments that will reveal relationships between channel proteins, and will give us insight into yet another type of order within the nervous system. □

A future for immobilized cell technology

from John F. Kennedy

ENZYMES are nature's catalysts and their ability to catalyse specific reactions under very mild conditions in a highly efficient manner has led to interest in their potential as industrial catalysts. However, enzymes are not always ideal for industrial applications as their instability usually prevents their use in, for example, organic solvents or at high temperatures. One of the major activities in the field of biotechnology over the past fifteen years has been the immobilization of enzymes to provide industrial catalysts that overcome some of these problems.

There have been numerous reports of laboratory-scale preparations of immobilized enzymes but the number of practical industrial systems using such systems has been very limited. Cheetham¹ and his colleagues at Tate and Lyle have just reported in *Nature* (299, 628) the production of a system in which the enzyme was immobilized without previous purification from its natural cell environment and the final catalyst performed under conditions in which the unbound enzyme would not, thereby allowing the development of a pilot-plant scale process which has been tuned to function with high efficiency in the high sugar (sucrose) concentrations required to aid the subsequent recovery of products by evaporation. An added bonus of the system is that the sugar concentration used prevents bacterial contamination of the biocatalyst.

The fundamentals of immobilized enzyme technology have been much discussed^{2,4}. But what of immobilized cell technology? Is this replacement to be regarded as a repugnant lowering of the standards of purity? Certainly not — immobilized cell in place of immobilized enzyme not only means a shorter and more economic route to producing the immobilized catalyst for a reaction, but opens up a whole vista of additional harnessing and manipulation of biological processes.

Besides asking the basic question: 'Why immobilize cells?', we need to ask ourselves again and again: 'What do we want to achieve by immobilizing cells?' Are we attempting to reproduce processes

normally conducted with free enzymes or free cells, or are we hoping, by using them, to find new processes hitherto undiscovered? In fact, what types of cells do we want to immobilize? Cells with the characteristics, reactions and molecular abilities necessary for the reaction we intend may not exist. Microbial cells⁵ are receiving almost all the attention at the moment simply because they are easily produced and readily multiply in conditions which can be achieved industrially. Plant cell immobilization has made a debut but given some availability of all cell types what will be our future choice, and why? Furthermore, should immobilized cells be 'dead' or alive and able to duplicate?

It seems to me that, because the immobilization of enzymes with retention of activity is a relatively new laboratory technique to which sophisticated chemistry and technology has been and must be addressed, and because the immobilization of cells with retention of cell life is even newer, we are all forgetting that many of the natural life processes depend on at least one of the following: (1) a phase change between substrate and cell — or tissue-bound enzyme (or other macromolecules), (2) interactions of solutes with cell surfaces, and (3) manifestation of chemical reactions by intact cells *per se* rather than release of soluble components from the cell for the purpose of participating in those chemical reactions. Thus, from the general viewpoint many lessons could be learnt from the chemistry of the natural processes of cell immobilization as in the building up of animal and plant tissues. However, such aspects of biological chemistry still need considerable attention and elucidation.

The methods which are used for the actual immobilization processes of cells⁶, as for enzyme immobilization, will necessarily be subjective and of individual choice. Therefore it will not be possible to suggest a universal means of immobilization. One can only say that as

the technology is developed, the search must continue for matrices which permit facile, secure immobilization with good interaction with substrate, and which conform in shape, size and so on to the use for which they are intended. At present all immobilizations are based on considered guesswork which results in a 'hit and miss' type of selection.

Choices must be made whether to immobilize cells by attachment to the surface of the matrix or by entrapment in the matrix. Relevant to this is the problem of molecular size of the substrate — entrapment may be more protective to the cell, but less approachable by the substrate. Entrapment may prevent multiplication of living cells.

In some cases a partial immobilization may be advantageous (as in some of our own work⁷ using hydrous metal oxides as matrix) in which cell density is increased by immobilization, but where the immobilization is reversible under certain conditions, allowing shedding of 'spent' cells as new ones grow. Questions must be asked as to what level of product contamination can be tolerated — after all, as soon as immobilized cells are mentioned, people tend to think of their being used for one enzyme reaction only and forget that cells are living systems which expose more than one type of molecule on the surface to the surrounding milieu and liberate a collection of molecular types and sizes to the surrounding milieu.

It is true that in the early stages of development of a new technology a detailed scrutiny from the viewpoint of economics can prove very discouraging. For immobilized cells to succeed in having a significant impact on certain aspects of our way of life they must either achieve something new and unavailable by other routes, or replace a (more) conventional process by doing it more cheaply. Thus cost considerations must take into account such factors as cell production, matrix type and form, cell immobilization conditions, reactor design, product purification and interactions of the cells with natural tissues. My prophecy for immobilized cell technology is that economic and advantageous processes can be devised but their general industrial adoption will need time and encouragement.

Cheetham and his colleagues have achieved the immobilization of microbial cells with retention of enzyme activity using a very mild technique which involves the formation of gel pellets of alginate in which the cell is in a resting state. The advantages of the use of this very topical support, obtained from seaweed, are not only cheapness, mechanical strength and non-toxicity, but also that the carbohydrate material provides an environment similar to the natural cell-wall polysaccharides produced by the microbial cells. The polysaccharides may assist in the stabilization of the cells.

The development of this system has been

John F. Kennedy is Director of the Research Laboratory for the Chemistry of Bioactive Carbohydrates and Proteins, University of Birmingham, PO Box 363, Birmingham B15 2TT.

taken a stage further than many similar methods by the construction of a pilot-plant scale process and has gone some way to dispel the criticisms of the Spinks report⁸ which concluded that Britain was again failing to take full commercial and industrial advantage of new technologies to which she has contributed so successfully at the primary research level. Not only has a system been reported which has scientific value, but by careful investigation, the industrial production of a naturally occurring disaccharide with great potential as a non-carcinogenic bulking agent for foodstuffs and pharmaceuticals can be realised with very few of the problems of purification and contamination associated with industrial

enzymic methods. It is hoped that this paper will stimulate further investigations into the industrial applications of immobilized cell technology. □

1. Cheetham, P.S.J., Imber, C.E. & Isherwood, J. *Nature* **299**, 628 (1982).
2. Wiseman, A. (ed.) *Handbook of Enzyme Biotechnology* (Ellis Horwood, Chichester, 1975).
3. Trevan, M.D. *Immobilized Enzymes. An Introduction and Applications in Biotechnology* (Wiley, London, 1980).
4. Kennedy, J.F. & Cabral, J.M.S. in *Solid Phase Biochemistry: Analytical and Synthetic Aspects* (ed. Scouten, W.H. (Wiley, New York, in the press).
5. Venkatsubramanian, K. (ed.) *Immobilized Microbial Cells Am. chem. Soc. Symp. Ser.* **106** (1979).
6. Kennedy, J.F. & Cabral, J.M.S. in *Applied Biochemistry and Bioengineering* (eds Wingard, L.B. & Chibata, I. (Academic New York, in the press).
7. Kennedy, J.F., Barker, S.A. & Humphreys, J.D. *Nature* **261**, 242 (1976).
8. *Biotechnology: Report of a Joint Working Party* (HMSO, London, 1980).

Immunoregulation by gamma-interferon?

from Teresa Basham and Thomas C. Merigan

THE potential of interferons as immunoregulatory as well as antiviral and antitumour agents is currently of great interest. With the advent of genetic engineering, pure preparations of biosynthetic interferons are now available for a precise examination of both their effectiveness in clinical trials and their mechanism of action. In this issue of *Nature* (p.833), David Wallach, Marc Fellous and Michel Revel describe a difference between pure recombinant gamma-interferon (IFN- γ) and the other interferons in the induction of the major histocompatibility antigens HLA-A,B,C which they suggest may be related to interferon's immunoregulatory function.

IFN- γ , a lymphokine derived from T lymphocytes, is one of three different types of interferons, each distinguished by protein sequence, antigenicity and physicochemical properties. The other two types are alpha-interferon (IFN- α) from leukocytes and beta-interferon (IFN- β) from fibroblasts. IFN- γ is closely linked to the immune response and can be induced by a wide variety of substances, including mitogens, bacterial and viral antigens, and tumour cells, while IFN- α and IFN- β are induced predominantly by viruses. In addition to antiviral activity, all three have been shown to have cytostatic and immunoregulatory functions, to modify cell structure, membranes and differentiation, and to induce the synthesis of several new proteins¹. The exact mechanism by which the different interferons exert these various activities is not

yet clear, and recent interest has focused on this problem.

Earlier work had shown that in the murine system, impure IFN- γ preparations increased the expression of the histocompatibility antigens (H-2) at doses 100–10,000 times lower than those required using IFN- α (ref. 2), and that IFN- γ is 250 times more potent in immune cell regulation than the other interferons^{3,4}. But these experiments used only partially purified interferons so it was possible the preparations contained other lymphokines. Using purified recombinant IFN- γ Wallach and his colleagues have now confirmed increased potency of IFN- γ in inducing HLA expression. Furthermore, they describe a differential response between the dose of recombinant IFN- γ required for the induction of HLA antigens and the enzyme (2'–5')oligoadenylate synthetase, which inhibits protein synthesis and is related to the antiviral activity of interferon. In contrast, IFN- α and IFN- β induced both of these cellular proteins at the same concentration. Thus Wallach and his associates suggest that the induction of the HLA antigens is not related to the antiviral activity of IFN- γ and that IFN- γ is the most important interferon in immune regulation.

Three recent papers showed that in a variety of cell types, interferon's effect on HLA expression occurs though a corresponding increase in the specific biosynthesis of these antigens^{5,6} and the mRNA pool for the HLA sequences^{6,7}. The optimum dose of interferon required for maximum HLA synthesis and expression varied from one cell line to another. For example, Basham *et al.* found that as low as 50 U ml⁻¹ of IFN- α were sufficient to induce maximum HLA biosynthesis and

expression in human melanoma cell lines sensitive to the antiviral activity of IFN- α but resistant to its antiproliferative action, whereas Burrone *et al.* used 2,000 U ml⁻¹ of IFN- α to induce a similar response in a T-cell line (Molt 4). Therefore, although it is intriguing that IFN- γ can induce HLA expression at a much lower concentration than that required for induction of (2'–5') oligoadenylate synthetase and antiviral activity, it is possible that any or all of these cell surface-related effects vary from one cell line to another, and many well characterized cell lines should be tested before a conclusion is drawn.

The function of the interferon-induced increases in HLA antigens is at present purely speculative but may well be related to cell–cell recognition. The most obvious role is in potentiating the cytotoxic T-cell response of a virally infected host for the elimination of HLA-matched infected cells. Wallach and co-workers suggest that the increase in HLA expression may also protect uninfected host cells from lysis by natural killer cells.

A claim for a role for IFN- γ , rather than the other interferons, in immune regulation would be much stronger if IFN- γ were shown to have a differential effect on the expression of the immune response antigens that are critical in antigen presentation by macrophages to lymphocyte subpopulations. Wallach *et al.* observed no increase in HLA-DR in a human B-cell line (Ramos). In contrast, using different cell types and impure IFN- γ , Sonnenfeld and colleagues² and, more recently, others⁸ found an increase in the human HLA-DR analogue, Ia, on murine cell surfaces. It is possible that Wallach *et al.* missed this effect of IFN- γ on HLA-DR because B-lymphoblastoid cell lines have a higher HLA-DR concentration on their surface initially.

More work on interferon's effects on antigen expression and regulation of the immune response will undoubtedly be done and should eventually clarify their contribution to the biological role of the interferons. □

1. Vilecek, J., Gresser, I. & Merigan, T.C. (eds) *Regulatory Functions of Interferon* (New York Academy of Sciences, New York, 1980).
2. Sonnenfeld, G., Meruelo, D., McDewitt, H.O. & Merigan, T.C. *Cell. Immun.* **57**, 427 (1981).
3. Sonnenfeld, G., Mandel, A.D. & Merigan, T.C. *Cell. Immun.* **34**, 193 (1977).
4. Virelizier, J.L., Chan, E.L. & Allison, A.C. *Clin. exp. Immun.* **30**, 299 (1977).
5. Basham, T.Y., Bourgeade, M.F., Creasey, A.A. & Merigan, T.C. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3265 (1982).
6. Burrone, O.R. & Milstein, C. *EMBO J.* **1**/3, 345 (1982).
7. Fellous, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3082 (1982).
8. Steeg, P.S., Moore, R.N., Johnson, H.M. & Oppenheim, J.J. *J. exp. Med.* (in the press).

ADDENDUM

The article 'Glomerular Challenge' at the Cretaceous-Tertiary boundary' was authored not just by three, but by all 18 of the DSDP leg 86 scientific staff.

Teresa Basham and Thomas C. Merigan are in the Division of Infectious Diseases, Department of Medicine, Stanford University School of Medicine, Stanford, California 94305.

Recent advances in optical bistability

from Y.R. Shen

IN recent years, optical bistability has been a subject of intense research in quantum electronics because of its potential usefulness in optical data processing and all-optical logic and computing systems¹. There has already been a conference on the topic², and at the XIIth International Quantum Electronics Conference held in Munich in June, two sessions were devoted to the subject.

In a system with optical bistability, the relationship of output to input is characterized by hysteresis loops; at a given input, the steady-state output can be either high or low depending on the operation path. These bistable characteristics are the basis of a binary switching element. Indeed, it was earlier predicted and demonstrated that an optical Fabry-Perot (F-P) interferometer filled with a nonlinear medium should exhibit bistability. The principle of the device, including the transient³ and steady-state behaviour¹ and the analogy with the phenomenon of phase transition⁴, is already fairly well understood. The more

recent advances on the subject are mainly in materials and technology.

A most significant advance in the survey of materials for nonlinear F-P interferometers is the discovery of the superiority of the GaAs-GaAlAs superlattice structure⁵. The future applications of optical bistability will require the development of small ($\sim 1 \mu\text{m}$) fast ($\sim \text{ps}$) low-operating-power ($\sim 1 \mu\text{W}$) room-temperature devices. Semiconductors with an excitonic — a positronium-like entity in the solid — line transition appear to be most promising because of their marked nonlinearity and the very low power required to achieve saturation in absorption.

Optical bistability has indeed been demonstrated in a $4.1 \mu\text{m}$ etalon of pure GaAs with an input laser beam at the exciton frequency⁶. Unfortunately, excitons in GaAs dissociate above 120K, while at low temperatures, the bistable switching time is long ($\sim \mu\text{s}$). But recently it has been found that GaAs-Ga_xAl_{1-x}As superlattices grown by molecular beam

epitaxy can have a larger exciton-binding energy than that of pure GaAs because of the quantum-well effect. As a result, excitons can exist in a superlattice even at room temperature. The power required for saturation of the exciton transition at room temperature is also three times less than that for pure GaAs.

These features point towards the possible construction of a small efficient optical bistable device. Indeed, on a $2 \mu\text{m}$ -thick etalon of GaAs ($335 \text{ \AA}$)-Ga_{0.73}Al_{0.27}As ($401 \text{ \AA}$) superlattice at room temperature, optical bistability can be observed with an $\sim 100 \text{ mW}$ laser at the exciton frequency. The switching time is around 20–40 ns. Further improvement can probably be made by using thinner layers in the superlattice structure.

Y.R. Shen is at the Department of Physics, University of California, Berkeley, and the Materials and Molecular Research Division, Lawrence Berkeley Laboratory, Berkeley, California 94720.



100 Years ago

LETTERS TO THE EDITOR

The Editor does not hold himself responsible for opinions expressed by his correspondents. Neither can he undertake to return or to correspond with the writers of, rejected manuscripts. No notice is taken of anonymous communications.

The Editor urgently requests correspondents to keep their letters as short as possible. The pressure on his space is so great that it is impossible otherwise to ensure the appearance even of communications containing interesting and novel facts.

The Comet

I ENCLOSE a drawing made this morning after a prolonged examination (with a binocular) of the end of the comet's tail. Should you think the peculiar features which I have endeavoured to portray of sufficient interest to reproduce, the drawing is at your service. It is difficult to indicate truly features of this kind without exaggeration, if they are to catch the eye at all; but I am sure the exaggeration is very slight. The tail would seem to be about to end rather suddenly and with a broad end, when, from near the middle, shoots out, at a slight inclination to the general direction of the tail, a cleanly-shaded wisp.

It is surely unusual for such decided features

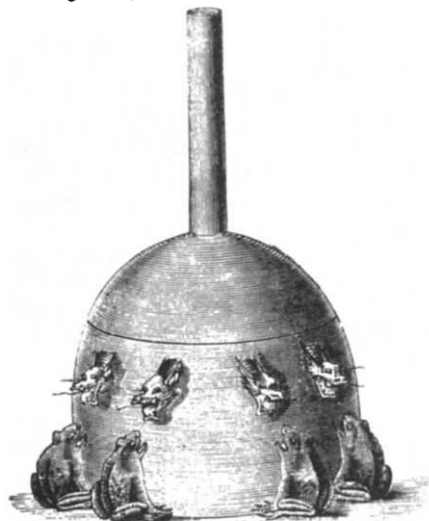


to present themselves at the very end of a comet's tail.

As a whole, the comet seems to have changed wonderfully little during the three weeks since I first saw it. Its change of place, also, is so moderate that, at this rate, there seems no reason why we should not see it for months yet. What if it should not vanish at all!

J. HERSHEL

Collingwood, October 23



In a Chinese history called "Gokanjo", we find the following: "In the first year of Yoka (A.D. 136) a Chinese called Chioko invented a seismometer. This instrument consists of a spherically formed copper vessel (Fig. 1), its diameter being 8 'shaku.' It is covered at its top. Its form resembles a wine bottle. Its outer part is ornamented with the figures of different kinds of birds and animals and old peculiar looking letters. In the inner part of this instrument a pillar is so placed that it can move in

eight directions. Also in the inside of this bottle there is an arrangement by which some record of an earthquake is made according to the movement of the pillar. On the outside of the bottle there are eight dragon heads, each of which contains a ball. Underneath these heads there are eight frogs, so placed that they appear to watch the dragon's face, so that they are ready to receive the ball if it should be dropped. All the arrangements which cause the pillars when it moves to knock the ball out of the dragon's mouth are well hidden in the bottle. When an earthquake occurs and the bottle is shaken, the dragon instantly drops the ball, and the frog which receives it vibrates vigorously. Any one watching this instrument can easily observe earthquakes. With this arrangement, although one dragon may drop a ball, it is not necessary for the other seven dragons to drop their balls unless the movement has been in all directions; thus one can easily tell the direction of an earthquake. Once upon a time a dragon dropped its ball without any earthquake, and the people therefore thought that this instrument was of no use, but after two or three days a notice came saying that an earthquake had taken place in Rosei. Hearing of this, those who did not believe about the use of this instrument began to believe in it again. After this ingenious instrument had been invented by Chioko, the Chinese Government wisely appointed a secretary to make observations on earthquakes."

We have here I think not only an account of an earthquake instrument which in principle is identical with many of our modern inventions, but the science has been conjoined with art. The record of the Chinese Government establishing a seismological bureau at a time when America was unknown, and half of Western Europe were living in the woods, is also interesting.

From *Nature* 26, 627; October 26, 1882.

Hitherto, the ray approximation has usually been used in the analysis of optical transmission in the nonlinear F-P interferometer. Because, on this view, different sections of the beam act independently in transmitting through the interferometer, if the beam has a gaussian profile, the switch-up and switch-down intensities would vary with radial position in the beam. But in reality, even in a medium with purely local response, different parts of the beam are coupled through diffraction, most markedly in F-P cavities with large Fresnel numbers. Moreover, the intensity-dependent transverse variation of the induced refractive index can distort the wavefront and lead to self-focusing or defocusing of the beam. Two papers have recently addressed the problem, and have shown by numerical calculation that strong coupling can result in a cooperative switch-up of practically the whole beam, and also in a much faster switching speed. The strong wavefront distortion can also give rise to an output with a multiple-ring pattern.

As a device with positive feedback, the stationary response of a bi-stable optical system may become unstable, leading to periodic oscillation, bifurcation and chaos — a subject of great current interest to many scientists in different disciplines.

Ikeda has recently spelled out, with some help of numerical calculations, the detailed conditions for the instability⁸. Although his theory uses the model of a ring cavity, the results should also apply, at least qualitatively, to the F-P interferometer. Instability, bifurcation and chaos have actually been observed in a F-P interferometer with electronic feed-back controlling the nonlinearity of the medium⁹. The experimental results are most useful as a test of the theory of chaos. Chaotic output from an all-optical nonlinear F-P interferometer would be even more interesting, but has not yet been realized. □

1. Gibbs, H.M., McCall, S.L. & Venkatesan, T.N.C. *Opt. Engng* 19, 463 (1980); Smith, S.D. & Miller, D.A.B. *New Sci.* 85, 554 (1980); Smith, P.W. & Tomlinson, W.J. *IEEE Spectrum* 18, 26 (1981).
2. *Optical Bistability* (eds Bowden, C.M., Cifan, M. & Robl, H.R.) (Plenum, New York, 1981); *IEEE J. Quantum Electron.* QE-17 (3) (1981).
3. Bischofberger, T. & Shen, Y.R. *Phys. Rev. A* 19, 1169 (1979).
4. Bonifacio, R. & Lugiato, L.A. in *Pattern Formation by Dynamic Systems and Pattern Recognition* (ed. Haken, H.) (Springer, Berlin, 1979).
5. Gibbs, H.M. *et al. Appl. Phys. Lett.* 41, 221 (1982); Miller, D.A.B. *et al. Pap. XIIth IQEC* (Munich, 1982).
6. Gibbs, H.M. *et al. Appl. Phys. Lett.* 35, 451 (1979).
7. Rosanov, N.N. & Semenov, V.E. *Opt. Commun.* 38, 435 (1981); Moloney, J.V. & Gibbs, H.M. *Phys. Rev. Lett.* 48, 1607 (1982).
8. Ikeda, K., Daido, H. & Akimoto, O. *Phys. Rev. Lett.* 45, 709 (1980); 48, 617 (1982).
9. Gibbs, H.M. *et al. Phys. Rev. Lett.* 46, 474 (1981).

Snail shells suggest past changes in British chalk vegetation

from Peter D. Moore

In view of the large proportion of southern Britain lying on chalky substrate and the amount of research effort expended on reconstructing the environmental history of the British Isles, it is surprising that so little is known about changes in the vegetation of the chalk over the past few thousand years. One reason is the scarcity of wetland sites in which lake sediments or peats could have accumulated and which could be suitable for palynological analysis. This is due in part to the pervasive nature of the substrate and intense land use. Such palynological data as do exist come mainly, in fact, from sites adjacent to, rather than on, the chalk, suggesting that these calcareous areas were once forested and lost their woodland cover during prehistoric times¹⁻³.

Analyses of plant remains from archaeological sites^{4,5} have provided detailed information concerning crop plants and weed species, and deposits in coombes at the foot of scarp slopes have been analysed for information on phases of erosion^{6,7}.

These sites have proved particularly informative when they contain charcoal fragments which can be carbon-14 dated — for example, the samples at Pitstone, Buckinghamshire dated at 3,910 BP by Valentine and Dalrymple⁸ indicated forest clearance in late Neolithic or early Bronze Age times.

Pollen does not normally survive in calcareous soils, but mollusc shells do and have been used extensively as environmental indicators⁹. For example, Evans¹⁰ has described the snail shells preserved in deposits adjacent to the Pitstone site and shown the replacement of woodland species by open, disturbed habitat species such as *Pomatias elegans*, thus confirming the pedological evidence for forest clearance.

A rather similar sequence has recently been investigated by Burleigh and Kerney¹¹ from the valley below a chalk escarpment at Brook, Kent, which is of particular interest because of its late-glacial sediments¹². As at the Pitstone site, the lower layers are rich in forest species, and then disturbance indicators, including *Pomatias elegans*, become gradually more abundant. In the layers where the transition takes place Neolithic flakes were found and a ¹⁴C date of 4,540 BP was

obtained from charcoal fragments, thus confirming the role of man in the deforestation process.

Burleigh and Kerney attempt to draw some palaeoclimatic conclusions from their snail shells. They find, for example, that the Neolithic *P. elegans* shells were generally larger than those of the present day, having a shell height of 12–18 mm (45 per cent of the population greater than 15 mm) in comparison with living snails which show a range of 11–15 mm. They state that shells higher than 15 mm are rather uncommon in Britain at the present time, although there are records of current populations with 7 per cent of individuals greater than 15 mm and occasional shells over 16 mm (ref. 13). Clearly, however, the Neolithic shells were generally larger than modern ones and Burleigh and Kerney associate the decline in shell size with falling temperature since the climatic optimum. *P. elegans* is a species with a southerly distribution in Britain and has contracted in its ranges since Atlantic times¹⁴.

They also note that 28 per cent of their fossil population of *Cepaea nemoralis* shells lacked banding, whereas only 4.3 per cent of modern *C. nemoralis* are found without bands. Currey and Cain¹⁵ considered that unbanded specimens are favoured by higher temperatures and Jones *et al.*¹⁵ have examined this idea closely and find that unbanded morphs of *C. nemoralis* are more frequent in the Mediterranean areas than in northern Europe. This can be explained in ecological/energetic terms, since unbanded morphs are unlikely to overheat in conditions of high insolation, whereas heavily banded morphs will absorb more radiation, which may be of advantage in cold conditions. On a local scale, other factors, such as predation, influence the balance of the polymorphism, so attempts to use such evidence in palaeoecological reconstructions must be cautious. Coupled with the evidence from *P. elegans*, however, Burleigh and Kerney are able to present a strong case that temperatures were warmer during the Neolithic than at present in this part of Britain.

A final point of interest in Burleigh and Kerney's studies is the use of snail shells as material for radio carbon dating. Normally such shells are avoided because old carbon

1. Thorley A. in *Guide to Sussex Excursions* (ed. Williams, R.B.G.) 47 (IBG Conf., 1971).
2. Godwin, H. *Veröff. Geobot. Inst. Rübel Zurich* 37, 83 (1962).
3. Brookes, A. Unpublished data from Wellingham, Sussex.
4. Allison, J. & Godwin, H. *New Phytol.* 48, 253 (1949).
5. Conolly, A.P. *New Phytol.* 40, 299 (1941).
6. Sparks, B.W. *Geol. Mag.* 89, 163 (1952).
7. Sparks, B.W. & Lewis, W.V. *Proc. geol. Ass.* 68, 26 (1957).
8. Valentine, K.W.G. & Dalrymple, J.B. *J. Soil Sci.* 27, 541 (1976).
9. Evans, J.G. *Land Snails in Archaeology* (Seminar, London, 1972).
10. Evans, J.G. *Proc. geol. Ass.* 77, 347 (1966).
11. Burleigh, R. & Kerney, M.P. *J. Archaeol. Sci.* 9, 29 (1982).
12. Kerney, M.P., Brown, E.H. & Chandler, T.J. *Phil. Trans. R. Soc.* B248, 135 (1964).
13. Moore, C.D. Unpublished data from Shoreham, Kent.
14. Kerney, M.P. *J. Conchology* 27, 359 (1972).
15. Jones, J.S., Leith, B.H. & Rawlings, P. A. *Rev. Ecol. System.* 8, 109 (1977).

Peter D. Moore is a Senior Lecturer in the Department of Plant Sciences, University of London King's College, 68 Half Moon Lane, London SE24 9JF.

may have become incorporated into them from the carbonate-rich environment either before or after the death of the animal. Having charcoal fragments, which are regarded as reliable material for dating, found in the same sediments as the shells provided an opportunity to examine their reliability in this respect. Dates from both

Pomatias and *Cepaea* shells were about 400–500 years older than those given by the charcoal. The older date is to be expected and an error of only about 10 per cent is surprisingly small. Although one must still be cautious, this is encouraging since for many studies this order of error is acceptable. □

Chromosome rearrangements in gonococcal pathogenicity

from J.R. Saunders

A STRIKING feature of *Neisseria gonorrhoeae*, the causative agent of gonorrhoea, is its ability to undergo substantial alteration in cell-surface characteristics during both natural infections and laboratory subculture. The most obvious of these changes is the reversible loss of ability to produce pili which confer adhesiveness and increased resistance to host defences^{1,2}. In a recent issue of *Cell*, Thomas Meyer and his colleagues³ show that the expression of pili is determined by rearrangements of the gonococcal genome. *N. gonorrhoeae* thus joins a number of pathogens, including *Salmonella typhimurium*^{4,5} and African trypanosomes^{6,7}, where moveable genetic elements affect the expression of surface antigens.

Single clinical isolates of *N. gonorrhoeae* exhibit several distinct colonial types when grown on the surface of clear agar media^{8–11}. Fresh isolates are invariably of the piliated (P^+) type. During subculture P^+ forms can give rise to non-piliated (P^-) forms and vice versa but after prolonged growth *in vitro* the P^- forms predominate. This suggests that the human host exerts considerable selective pressure on the gonococcus to produce pili.

Gonococcal colonies also vary in opacity, being either opaque (O^+) or translucent (O^-). Opacity is determined by the presence of a family of outer membrane proteins (protein II) that are associated with alterations in the adhesive and antigenic properties of the bacteria^{9,10}. Variation between O^+ and O^- colonial forms occurs reversibly in a similar way to pilus expression.

There is considerable variation in the physical and antigenic properties of pili in different strains of *N. gonorrhoeae*^{12,13}, and even in derivatives of the same strain^{16–18}. For example, variants of *N. gonorrhoeae* strain P9 produce four types of pili— α , β , γ and δ —of subunit molecular weights 19,500, 20,500, 21,000 and 18,500 respectively^{17,18}. The four pilus variants show little antigenic cross-

reactivity and have differential adhesiveness for a number of animal cell types¹⁹. However, they possess similar amino acid compositions and have a number of major peptides in common²⁰. It seems that the first 100 amino acids on the N-terminal end of pilus subunits are conserved among serologically distinct gonococcal pili whereas the C-terminal ends of the proteins are variable^{14,15}.

Variation in the type of pilus and in outer membrane composition may be important in permitting the gonococcus to colonize different sites on the human host during the course of natural infections. Inter- and intra-strain variability would also help a pathogen to avoid host defences.

Analysis of the genetic control of pilus and outer membrane expression in the gonococcus has been hampered by the relative difficulty of performing conventional genetics with a pathogen that exhibits such variation in natural phenotype. Nonetheless, the pilus gene from strain MS11 (ref.3) and the α -pilin gene from strain P9 (ref.21) have recently been cloned in *Escherichia coli*. The gonococcal pilin genes are expressed in *E. coli*, but no mature pili can be detected on the surface of this host^{3,21}. This is perhaps not surprising since several genes are known to be required for pilus production and assembly of other types of pili, for example K88 (ref.22). Even if all the required genes were cloned it is still possible that *E. coli* would be unable to process and insert the pilin molecules into the cell envelope to form an intact pilus.

This cloning of pilus genes has provided probes for the investigation of the molecular switching events involved in alterations in virulence of the gonococcus. Thus Meyer *et al.*³ have used the cloned pilin gene from strain MS11 as a probe in Southern hybridizations of restriction endonuclease-cleaved chromosomal DNA from various derivatives of this strain and shown that the conversion from piliated (P^+O^+ or P^+O^-) to non-piliated (P^-O^+ or P^-O^-) form is accompanied by a chromosomal rearrangement (presumably in the pilus region) of the gonococcal genome. Apparently, the same rearrangement is found after conversion

from P^+O^- to P^+O^+ . It is not yet clear, however, whether this rearrangement actually represents control of both pilus and protein II synthesis by the same genetic element or merely results from a coincidental change in the type of pilus synthesized in this particular experiment. Certainly piliation and opacity characteristics can vary independently^{23,24}. Confirmation or denial of a common controlling element will depend on the use, when it becomes available, of cloned DNA encoding protein II. The hybridizations performed by Meyer *et al.* also indicate that the chromosome of *N. gonorrhoeae* MS11 contains multiple copies of the pilin gene or at least of the constant region (encoding the first 100 amino acids) of that determinant, reflecting the importance of pili to the organism.

The rearrangements of the gonococcal chromosome appear similar to those mediating antigenic phase variation in the flagella of *Salmonella typhimurium*^{4,5}. In this case, a 970-base pair DNA segment containing a promoter for the flagella gene can invert reversibly by site-specific recombination at the ends of the segment. Movement of non-expressing surface antigen genes to new expressing sites also explains antigenic variability in trypanosomes and the ways in which these parasites can evade the host immune response^{6,7}.

The gonococcus apparently possesses both an on-off switch for pilus production and a means of altering the protein composition of the pili it produces. A similar switching system seems to control protein II production. The simplest explanation for these changes would

1. Buchanan, T.M. in *The Gonococcus* (ed. Roberts, R.B.) 255 (Wiley, New York, 1977).
2. Ofek, I., Beachy, E.H. & Bisno, A.L. *J. infect. Dis.* **129**, 310 (1974).
3. Meyer, T.M., Mlawer, N. & So, M. *Cell* **30**, 45 (1982).
4. Zieg, J., Silverman, M., Hilmen, M. & Simon, M. *Science* **196**, 170 (1977).
5. Zieg, J. & Simon, M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4196 (1980).
6. Williams, R.O., Young, J.R. & Majewa, P.A.O. *Nature* **282**, 847 (1979).
7. Kocjmakers, J.H.J. *et al. Nature* **284**, 78 (1980).
8. Kellogg, D.S. *et al. J. Bact.* **85**, 1274 (1963).
9. Swanson, J. *Infect. Immun.* **19**, 320 (1978).
10. Lambden, P.R., Heckels, J.E., James, L.T. & Watt, P.J. *Gen. Microbiol.* **114**, 305 (1979).
11. Penn, C.W., Veale, D.R. & Smith, H. *J. gen. Microbiol.* **100**, 147, (1977).
12. Brinton, C.C. *et al. in Immunobiology of Neisseria gonorrhoeae* (ed. Brooks, G.F.) 155 (American Society for General Microbiology, Washington).
13. Buchanan, T.M. *J. exp. Med.* **141**, 1470 (1975).
14. Hermodson, M.A., Chen, K.C.S. & Buchanan, T.M. *Biochemistry* **17**, 442 (1978).
15. Chen, K.C.S. Cited in ref. 3.
16. Salit, I.E., Blake, M. & Gotschlich, E.C. *J. exp. Med.* **151**, 716 (1980).
17. Lambden, P.R., Robertson, J.N. & Watt, P.J. *J. Bact.* **141**, 393 (1980).
18. Lambden, P.R., Heckels, J.E., McBride, H. & Watt, P.J. *FEMS microbiol. Lett.* **10**, 339 (1981).
19. Heckels, J.E. in *Microbiology 1982* (ed. Schlessinger, J.) (American Society for Microbiology, Washington, in the press).
20. Lambden, P.R. *J. gen. Microbiol.* **128**, 2105 (1982).
21. Heighway, J., Graeme-Cook, K.G., Heckels, J.E. & Saunders, J.R. (in preparation).
22. Mooi, F.R., De Graaf, F.K. & van Embden, J.D.A. *Nucleic Acids Res.* **3**, 849 (1979).
23. Meyer, L.W. *Infect. Immun.* **37**, 481 (1982).
24. Norlander, L., Davies, J., Norqvist, A. & Normark, S. *J. Bact.* **138**, 762 (1979).

J.R. Saunders is in the Department of Microbiology, The University of Liverpool, PO Box 147, Liverpool L69 3BX.

involve some sort of transposable genetic element capable on the one hand of allowing or preventing pilus synthesis and on the other of altering the quality of the pilin produced (presumably by altering only the C-terminal end of the pilus protein). It is intriguing that a set of equal-sized (160 base pair) *Hae*II restriction fragments can be detected both within and adjacent to the pilus-coding region of the MS11 genome. Whether these are part of some invertible sequence controlling pilus

expression remains to be seen. It is also unclear whether post-translational modification of envelope components has any role in mediating the changes in gonococcal surface properties.

Nevertheless, the observation that chromosome rearrangements can affect virulence is a very important step towards understanding the extremely complex genetic and phenotypic interactions between this important pathogen and its human host. □

Bacterial toxins and cyclic AMP

from Simon van Heyningen

It has been known for several years that cholera toxin and the very similar protein toxin of some strains of *Escherichia coli* enter intoxicated cells and activate the intracellular adenylate cyclase that catalyses the formation of cyclic AMP from ATP. In cholera patients, this increase in cyclic AMP in gut cells is the chief cause of the massive diarrhoea that is characteristic of the disease. Several recent papers have shown that other toxins also increase the cyclic AMP concentration. In view of the demonstrated connection between adenylate cyclase and cholera, it is surprising that, until recently, little attention had been given to the possibility that the soluble adenylate cyclases secreted by a number of bacteria play a role in virulence. (These cyclases differ from the mammalian enzymes in that they are more active, soluble, heat stable and are not usually regulated by guanine nucleotides.)

A picture is now emerging, however, as a result of recent work, particularly that by Leppla¹, on the toxins of *Bacillus anthracis* which causes anthrax. Three different proteins produced by this organism can be purified: protective antigen (PA or factor II), lethal factor (LF or factor III) and oedema factor (EF factor I). None of these has much biological effect by itself, but combinations of PA with either of the others do: PA and EF produce a skin oedema, and PA and LF are lethal.

Leppla guessed that EF might have a similar action to that of cholera toxin because it has similar effects on vascular permeability in rabbit skin and on the morphology of Chinese Hamster Ovary (CHO) cells. Accordingly, he showed that PA and EF together greatly increased the cyclic AMP concentration in CHO and other cultured cells — however this was not due to increased activity of the adenylate cyclase in the cells, but to enzymatic

activity of EF itself. When EF was tested *in vitro*, it was found to be a highly active cyclase, but only in the presence of a CHO cell lysate. The lysate could be replaced by calmodulin (known to be an activator of many adenylate cyclases) and the activity of the enzyme varied with calcium ion concentration in a manner characteristic of calmodulin-dependent systems. In whole cells, EF was active only in the presence of PA which presumably functions in some way to aid EF entry into the cell. (It must also have this function with the other protein toxin, LF, which is not a cyclase.) EF was thus the first example of a bacterial cyclase that enters eukaryotic cells and is activated by a eukaryotic protein. The relevance of all this to anthrax is not clear — the mixture of PA and LF has generally been thought to be the lethal agent, and they may mask the effects of adenylate cyclase during the disease.

Bacillus anthracis is not the only pathogen that secretes adenylate cyclase. *Bordetella pertussis*, responsible for whooping cough, also produces a highly active adenylate cyclase that is heat stable and stimulated by calmodulin². Confer and Eaton³ have shown that this enzyme can be taken up by phagocytic cells and generate cyclic AMP. Human neutrophils incubated with extracts from *Bordetella* prepared by urea extraction followed by dialysis showed a striking increase in intracellular cyclic AMP concentration, and the cyclase activity in the cells was heat stable, showing that it must have come from the bacterium. How it got into the cells is not clear — purified bacterial cyclase is very active but only with broken cells. A temperature-dependent process is involved, as well as other factors secreted by *Bordetella* (perhaps analogous to the PA of anthrax toxin).

What is the function of the increase in cyclic AMP? It is possible it is used by the invading bacteria to suppress many of the normal functions of phagocytes, including bacterial killing, and so allow the bacteria

to survive longer in a hostile environment. It may also explain the impaired host defence, for example the absence of fever and frequent secondary bacterial pneumonias, found in whooping cough.

Toxin production by *Bordetella pertussis* is complicated and confusing. As well as the extracellular cyclase, the bacterium secretes as many as twenty other antigens. One of these, known as 'pertussis toxin', or islet-activating protein, because there is good evidence that it causes the main harmful effects of whooping cough, has many biological effects, such as potentiation of the immune response, induction of lymphocytosis and sensitization to histamine, which may also be due to an increase in cyclic AMP. However the toxin itself is not a cyclase — it works, like cholera toxin, by activating the cyclase of the intoxicated cell. ATP and NAD⁺ are required in broken cell preparations and, like cholera and diphtheria toxins, pertussis toxin catalyses the transfer of ADP-ribose from NAD⁺ to a receptor protein⁴. This receptor protein has a molecular weight of 41,000 and is presumably part of the adenylate cyclase complex, but it is not the same protein as the one ADP-ribosylated by cholera toxin. Thus *Bordetella pertussis* secretes both an adenylate cyclase of its own and a protein that activates another cyclase. Perhaps the 'toxin' is responsible for the overt symptoms of disease and the cyclase for the ability of the bacteria to survive.

It is remarkable that bacteria can produce proteins that have such similar properties to eukaryotic cyclases and that can be activated by a eukaryotic protein — calmodulin — which is not known to occur in bacteria. It suggests the proteins might have originated in some eukaryotic cyclase which the bacteria have picked up during the course of their evolution.

The same suggestion has been made of the toxins that catalyse ADP-ribosylation. Some recent work suggests that the hormone thyrotropin can increase ADP-ribosylation of several proteins, including the substrate of cholera toxin, in thyroid cell membranes at the same time as adenylate cyclase activity⁵. It seems that the hormone does not catalyse the reaction directly, but increases the activity of endogenous ADP-ribosyl transferases which are known to be present and are presumably part of the normal control process. Cholera, pertussis and other opportunistic toxins are thus subverting this process to their own ends. Since cholera and pertussis toxins ADP-ribosylated different proteins, the system must be a very complicated one. □

Simon van Heyningen is at the Department of Biochemistry, University of Edinburgh Medical School, Edinburgh EH8 9XD.

1. Leppla, S.H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3162 (1982).
2. Greenlee, D.V., Andreasen, T.J. & Storm, D.R. *Biochemistry* **21**, 2759 (1982).
3. Confer, D.L. & Eaton, J.W. *Science* **217**, 948 (1982).
4. Katada, T. & Ui, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3129 (1982); *J. biol. Chem.* **257**, 7219 (1982).
5. Vitti, P., De Wolf, M.J.S., Acquaviva, A.M., Epstein, M. & Kohn, L.D. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1525 (1982).

REVIEW ARTICLE

Hot galactic gas and narrow line quasar absorption systems

T. W. Hartquist & M. A. J. Snijders

Department of Physics and Astronomy, University College London, London WC1E 6BT, UK

Observations of the hot interstellar gas ($T \gg 10^4$ K) are numerous and imply that the temperatures of substantial amounts of the disk gas to be $2\text{--}3 \times 10^5$ K and of the halo gas to be 8×10^4 K, but the theoretical models to explain the properties of both the disk and the halo hot gas seem to be in need of some refinement. Observations of the quasar narrow line absorption features suggest that they arise in gas with properties very similar to those of the galactic coronal gas.

WITHIN the past decade the Copernicus and IUE satellites have made possible the detailed study of the hot interstellar gas ($T \gg 10^4$ K) in the galactic disk and in the galactic halo. Additional observations have been made, using rockets, satellites, and manned spacecraft, of soft X-ray and EUV emission in the Galaxy. It is generally believed that some of these soft X rays are emitted by the hot interstellar gas. We now review the UV and soft X-ray observations and argue that the UV data imply that the hot interstellar gas has a very low thermal pressure whereas the soft X-ray data imply a very high pressure. We discuss existing models of the hot disk interstellar gas concluding that the UV and soft X-ray observations are difficult to explain.

The temperature of the observed halo corona gas appears to be considerably lower than that in the disk and to be remarkably uniform. Although models of the halo gas have been published, only recently have any of them been designed to explain the uniform temperature of the halo gas.

The most important consequence of the galactic observations is that they seem to indicate that the temperature of the hot disk gas is $\sim 2\text{--}3 \times 10^5$ K and the temperature of the hot halo gas is $\sim 8 \times 10^4$ K. These temperatures are far lower than the value of 10^6 K which has been accepted since 1956 when Spitzer¹ reported on the galactic corona which he postulated to explain the stability of high latitude clouds. Note that Spitzer presented isothermal models of the corona for temperatures ranging over two orders of magnitude and that he discussed the diagnostic utility of absorption features of ions which would be abundant at temperatures below 8×10^4 K to $\sim 10^6$ K. The value 10^6 K seems to have become established because Spitzer quoted the sound speed, the density, and the total mass of his model for 10^6 K apparently just to give some idea of the magnitudes of the physical quantities he was discussing: Spitzer did not argue in any way that $T \approx 10^6$ K.

Many of the ions abundant in the galactic coronal gas are abundant in the narrow line quasar absorption systems. We compare the observational properties of the galactic corona with these systems.

Observations of the hot disk gas

UV absorption observations of several moderately ionized species yield information about the physical conditions in the corona. Table 1 lists several important species, the wavelengths of their absorption features, and the temperatures at which, in the steady state coronal approximation, they reach their maximum concentrations. In Fig. 1 the fractional abundance of these species is plotted as a function of temperature in the steady state coronal approximation. Al^{2+} lies in gas which is not in ionization equilibrium² and is not included in Fig. 1. The silicon results are from Baliunas and Butler³ who were the first to include several important charge transfer processes, and the

C^+ , N^{4+} , and O^{5+} results are those of Shapiro and Moore⁴. We adopted the cosmic abundances given by Allen⁵.

Interstellar O VI was detected⁶⁻⁸ in the directions of several early type stars indicating the presence of interstellar gas at temperatures in excess of 10^5 K. Castor *et al.*⁹ and Weaver *et al.*¹⁰ suggested that the O VI absorption could arise in conduction fronts at the interfaces between shocked interstellar material and stellar winds driving the supersonic expansion of cavities or bubbles around the early type stars towards which the observations were made. However, Jenkins¹¹ has shown that the O VI is not an artefact of the wind-interstellar medium interaction by performing a statistical analysis to find that the column densities of O VI observed against stars of varying distances is in accord with the existence of randomly distributed cells of hot gas having O VI column densities of 10^{13} cm^{-2} . The average spatial number density of O VI is $\sim 2.8 \times 10^{-8} \text{ cm}^{-3}$.

York^{7,12} reported the first detection of interstellar N V, which arises in the same regions as O VI. Towards β Cen and α Ori he found that the N V to O VI abundance ratio was ~ 0.1 while towards HD28497 and μ Col the upper bounds on the ratio were roughly this value. Figure 1 shows that the temperature of the gas would have to be $\leq 3 \times 10^5$ K to be consistent with his results. Smith and Hartquist¹³ published observations of N V in the galactic disk gas towards eight Wolf-Rayet stars which, because of their extremely dense winds, have exceptionally strong broad emission features against which to detect interstellar absorption. Because some of the stars in their study are as much as 3 kpc away the derived average number density

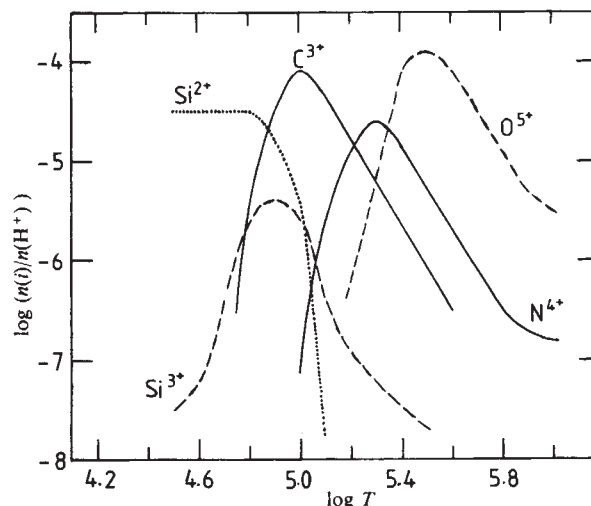


Fig. 1 The fractional abundance of several ionized species as a function of temperature.

of $6-7 \times 10^{-9} \text{ cm}^{-3}$ results from sampling the physical conditions in a large amount of gas. Their results and that of Jenkins¹¹ for the average number density of O VI give $\log T \approx 5.4$. A temperature in this range is also required to explain the narrowness of many of the O VI features¹¹ discussed above. Bruhweiler *et al.*¹⁴ found N V along the line of sight to AO Cas, one of the most distant stars in their sample; the average line of sight density of N^{4+} was comparable with that found by Smith and Hartquist. Bruhweiler *et al.* obtained upper limits for N V in the other directions by assuming that their detection limit was 50 mÅ in all directions. If this limit is correct for the other three most distant stars in their sample the gas in those directions would have to have temperatures higher than that in the AO Cas direction. Black *et al.*¹⁵ and Cowie *et al.*¹⁶ saw N V absorption in about a quarter of the directions in which they looked. The results in these directions too are in harmony with those reported by Smith and Hartquist and the upper limits in the other directions are not very stringent.

From the N V to O VI ratio one can conclude that the temperature of the hot interstellar gas lies in the range $\log T = 5.4 \pm 0.1$. Calculations by Shapiro and Moore⁴ of the N V to O VI ratio in cooling gas indicate that this conclusion is independent of the history of the gas. At $\log T = 5.4$, $\log(n(\text{O}^{5+})/n(\text{H}^+)) = -4.3$ and the product of the hot phase's thermal pressure and its filling factor is $\sim 30 k \text{ erg cm}^{-3}$ where k is Boltzmann's constant. Even if the filling factor of the hot gas is as low as 0.2 (ref. 11) its thermal pressure is an order of magnitude below that of the 'standard' H I clouds¹⁷. Either only H I clouds in high-pressure regions are detected or the H I clouds are confined by some other source of pressure. Turbulent pressure is unlikely to be great enough in the hot phase as the turbulence would have to be supersonic and would dissipate rapidly. Perhaps H I clouds are confined by magnetic pressure¹⁸. As is known from attempts to confine laboratory plasmas¹⁹ the clouds can be confined by field pressure only if the configuration is not subject to the interchange instability²⁰. If some H I clouds are contained by magnetic pressure, they may form initially by thermal instability and collapse along the field lines to a density of $\sim 1 \text{ cm}^{-3}$ and a temperature of around 10^2 K . Draine²¹ has argued that the thermal pressure can be enhanced relative to the magnetic pressure in interstellar material with these physical conditions by the passage of a hydromagnetic shock in which ambipolar diffusion (the motion of the ions, which couple directly to the magnetic field, relative to the neutral particles which are not well coupled to the ions and the magnetic field) occurs.

When O VI lines were first detected it was hoped that measurements of C IV and Si IV might help diagnose the conditions in the hot gas^{6,7}, but much of the C IV and Si IV detected in the disk^{14-16,22} is apparently formed by photoionization in normal H II regions^{15,16}. As shown in Fig. 1 the concentrations of interstellar C IV and Si IV will be far lower than that of N V at temperatures above $2 \times 10^5 \text{ K}$. Black *et al.*¹⁵ and Laurent *et al.*²³ reported the detection of collisionally ionized gas with C IV and Si IV towards the Carina Nebula and Smith *et al.*²² reported similar detections in several other directions in which the intervening gas appears to be disturbed. Cowie *et al.*¹⁶ probably found collisionally ionized gas towards two distant disk stars with an average spatial density of $2 \times 10^{-9} \text{ cm}^{-3}$ for C^{3+} . If C^{3+} has such a large spatial density, it is roughly one-tenth as abundant as O^{5+} . Figure 1 shows that these two abundances can be in that ratio only in a very narrow temperature range around $T = 10^{5.55} \text{ K}$ if the gas is in steady state collisional equilibrium.

The existence of hot interstellar gas has been invoked to explain the origin of the soft X-ray and EUV background^{24-30,93}. Rosner *et al.*²⁶ have argued that much of the flux between 0.28 and 1 keV is produced by M-dwarf stars, and Levine *et al.*²⁷ list known supernova remnants which are localized sources contributing substantially to the soft X-ray and EUV backgrounds. The 0.15–0.28 keV background apparently is not of stellar origin²⁶. From observations between 90 and 140 eV

Levine *et al.*⁹³ concluded that the gas producing soft X rays has a temperature in excess of $6 \times 10^5 \text{ K}$, while Cash *et al.*²⁸ used similar arguments to conclude that the gas must be at temperatures below $3 \times 10^5 \text{ K}$ to produce the high flux at the very longest wavelengths, first observed by Yentis *et al.*²⁹. Interpretation of EUV data obtained from the Apollo–Soyuz mission³⁰ seems to confirm the result of Cash *et al.*²⁸. A temperature of $3 \times 10^5 \text{ K}$ is in harmony with the results of the N V and O VI comparison and the narrow widths of the O VI features, but the high emission measure required by the EUV data restricts the number density of $3 \times 10^5 \text{ K}$ gas to exceeding 10^{-2} cm^{-3} , much larger than the one obtained from the O VI observations.

A 1,350–1,550 Å background at high galactic latitudes may exist³¹. If so its origin may be due to the scattering of stellar radiation by dust grains in high latitude clouds³². The rapid growth of extinction³³⁻³⁶ at 1,000 Å suggests that many grains in diffuse clouds have radii of $\leq 10^{-6} \text{ cm}$ (ref. 17). Possibly, the extinction due to dust is very great at $\lambda \sim 100 \text{ Å}$ and is due to scattering. The diffuse EUV and very soft X-ray backgrounds then could originate in a few localized sources, such as supernova remnants, with high densities. Their radiation would scatter on the intervening clouds. Overbeck³⁷ and Hayakawa³⁸ have discussed the origin of soft X-ray haloes around point sources from such scattering. If the extinction and albedo at 100 Å are high several scatterings will occur before significant numbers of soft X rays are absorbed. Indeed, scatterings may prevent soft X rays from penetrating far enough into neutral clouds for them to suffer much attenuation even though their source is very far away.

Observations of the halo gas

Several moderately ionized species have been detected in absorption towards bright stars in the Magellanic clouds^{39,40}, towards bright stars in the galactic halo^{41,42}, and towards 3C273⁴³. C IV and Si IV were detected towards virtually all sources $> 1 \text{ kpc}$ above the disk, and these lines were significantly stronger than those seen towards galactic disk stars. Further the C IV to Si IV ratio in the galactic halo is not typical of that for photoionized gas in H II regions⁴². Hence, it is generally accepted that C^{3+} and Si^{3+} are very abundant in the hot halo interstellar gas. Halo N V has been detected in only one line of sight towards the Small Magellanic Cloud⁴⁴ and towards two halo stars⁴². The latter two lines of sight pass through considerable amounts of disk material as the two background stars are not at very high latitudes and the N V features in those directions may arise in the hot disk gas. The feature seen towards the Small Magellanic Cloud star may arise in an intervening supernova remnant, a few of which may exist in the halo if star formation in high latitude, intermediate and high velocity clouds occur⁴⁵. Typically the observed equivalent widths of the stronger components of the doublet absorption features were 200 mÅ for Si IV and 400 mÅ for C IV. The detections of N V were at the 50-mÅ level.

If the gas, in which C IV and Si IV is seen, is in steady-state collisional equilibrium and if its abundances of carbon and silicon are solar its temperature is $\sim 8 \times 10^4 \text{ K}$. The observed ratios of the features are very uniform and the total possible spread in temperature in the observed regions is only $\sim 1 \times 10^4 \text{ K}$ (ref. 42). At $8 \times 10^4 \text{ K}$, N^{4+} is very scarce; the upper bounds in most directions limit the halo gas temperature to being $< 1.5 \times$

Table 1 UV absorption features

Ion	T (max concentration) (K)	Wavelength (Å)
Al III	$10^{4.6}$	1,862.79, 1,854.72
Si III	$10^{4.3-10^{4.8}}$	1,206.51
Si IV	$10^{4.9}$	1,402.77, 1,393.76
C IV	$10^{5.0}$	1,550.77, 1,548.20
N V	$10^{5.3}$	1,242.80, 1,238.82
O VI	$10^{5.5}$	1,037.63, 1,031.94

Table 2 Total column densities along a line of sight perpendicular to the galactic disk

Ion	Disk		Refs	Halo		Refs	Total $N(10^{13} \text{ cm}^{-2})$
	$n(10^{-9} \text{ cm}^{-3})$	$N(10^{13} \text{ cm}^2)$		$n(10^{-9} \text{ cm}^{-3})$	$N(10^{13} \text{ cm}^{-3})$		
C^{3+}	≤ 2.0	0.37	16, 23	8.5	15.8	39, 41, 42	16.1
Si^{3+}	0.5	0.09	13	1.9	3.5	39, 42	3.6
N^{4+}	5.0	0.86	13	< 1.5	< 2.78	39, 42	< 3.6
O^{5+}	28.0	5.20	11	?			> 5.2

10^5 K. York⁴⁶ has discussed the possibility that the haloes Si^{3+} and C^{3+} are formed by the photoionization of lower ionized species in warm interstellar gas by an intergalactic UV background produced by quasars and other active galaxies. However, the lowest energy photons will be absorbed relatively near the top of the corona. There, the fractional ionization will be high and the charge transfer ionization³ of Si^+ to form Si^{2+} will be rapid and Si^{3+} will be formed by the photoionization of Si^{2+} . Only a bit lower in the halo the fractional ionization will have dropped and the charge transfer recombination of Si^{2+} will be rapid and Si^{3+} will not be formed⁴⁷. The production of Si^{3+} and C^{3+} by photoionization cannot explain their high abundances over a substantial range of heights above the disk.

Estimates for the vertical distribution of the hot halo gas were obtained by observing stars at different heights above the plane^{41,42} and by observing the velocity profile of the features and assuming that the halo gas co-rotates with the disk^{39,40}. Bromage *et al.*⁴¹ and Savage and de Boer^{39,40} found that the scale height is about 2 or 3 kpc which compares with the scale height of 1.4 kpc originally given by Spitzer for 10^5 K gas. Pettini and West⁴², who studied the most extensive sample of objects, reported that while most of the C IV and Si IV absorption occurs within 3 kpc of the plane the number densities of C^{3+} and Si^{3+} may increase between 0.5 and 3 kpc. If so a static plane stratified model of the halo gas may be inappropriate.

The total column density of gas at 8×10^4 K along a line of sight perpendicular to the disk is $\sim 10^{19} \text{ cm}^{-2}$. If the scale height is 3 kpc its number density at the edge of the disk is 10^{-3} cm^{-3} . A number density of 10^{-3} cm^{-3} , a temperature of 10^5 K, and a filling factor of about 1 correspond to a pressure comparable with that which we derived for the hot disk gas from the N V and O VI data. If f is the true filling factor of the 8×10^4 K gas it will radiate about $0.1 f^{-1}$ of the total galactic supernova energy⁴⁸. If the 8×10^4 K gas were embedded in hotter gas the dissipation of energy in the 8×10^4 K gas would have to be almost unrealistically efficient.

Si III and Al III have also been detected in the hot halo gas⁴⁰. The presence of Si III is expected if the halo gas is at 8×10^4 K, but the Al III feature should be unobservable if all of the gas is at such temperatures. Al^{2+} can be abundant in cooling gas if its initial temperature is $< 5 \times 10^4$ K (ref. 2). Al^{2+} in the halo may form in gas cooler than 5×10^4 K which is above the 8×10^4 K gas and is unstable to cloud formation. If this is its origin the presence of Al^{2+} puts restrictions on the coronal fountain models discussed later.

Many absorption features arising in much cooler halo gas also were seen in the UV⁴⁰. Cool predominately neutral high velocity clouds exist at high galactic latitudes⁴⁹. While they undoubtedly lie in the halo little is known about their location or kinematics. If they are moving through 8×10^4 K halo gas shock heating of the ambient gas occurs. Soft X rays emitted during the subsequent cooling of the gas can serve as diagnostics of their kinematics⁵⁰. The formation of the clouds and the explanation of their kinematics present important challenges to any theory of the hot halo gas.

Twenty-one centimetre studies⁵¹ revealed that warm neutral atomic hydrogen with $T = 1-2 \times 10^3$ K exists at high galactic latitudes. It has been suggested⁵² that such gas can arise at the interface between the more distant edge of molecular cloud and the hot gas. In this model the stability of the gas at $1-2 \times 10^3$ K is due to the presence of molecular hydrogen, which is an efficient coolant.

Theories of coronal heating

Bottcher *et al.*⁵³ made the earliest attempt to construct a self-consistent model of the interstellar medium in the disk taking account of the ionization and heating due to supernova remnants. Since the discovery of O VI absorption⁶ in the interstellar medium other models in which supernova remnants have a role in regulating the balance of the interstellar medium have appeared. Some of the general ideas used in these models in which remnants either tunnel hot channels in the interstellar medium⁵⁴ or overlap before cooling and fill most of the interstellar space⁵⁵ are probably applicable. However, the models face severe difficulties. For example, the model of McKee and Ostriker⁵⁵ has gas pressures too low to explain the origin of the EUV background³⁰. It predicts too much O^{5+} which the authors claim could be hidden if turbulent motions broaden the lines; however, the turbulence would have to be supersonic and would dissipate rapidly and the observed lines are narrow. The model predicts that N V is undetectable^{13,16}. The soft X-ray-EUV, the O VI, and the N V results are the three primary observations of hot interstellar disk gas that a successful model of the interstellar medium must explain.

The hot gas in the Galaxy may be divided into two types. One is that which is close to the plane and near stars among which supernovae do occur. O^{5+} has a density distribution with a scale height of 300 pc (ref. 11) and probably exists in this first type of gas. Gas further from the plane is probably heated by energy generated in the disk by sources including supernova remnants, but the energy is transmitted above the plane by mass flow, cosmic rays, or hydromagnetic waves.

Weisheit and Collins⁵⁶ who calculated the ionization balance for the hot halo gas reviewed the possible heating mechanisms suggested before early 1976. Heating by soft X-ray induced photoionization or by infalling extragalactic material⁵⁷ is insufficient. The conversion of cosmic ray energy by the generation of Alfvén waves which decay rapidly by nonlinear wave-wave interactions⁵⁸⁻⁶⁰ would be a sufficient source of heat. The observation that the detected coronal gas is at 8×10^4 K may prove valuable in constraining models of galactic cosmic ray propagation.

Chevalier and Gardner⁶¹ and Cox and Smith⁵⁴ suggested that some of the hot gas in the plane escapes and rises into the halo. Chevalier and Oegerle⁶² have argued that such flows probably could not lead to the production of both a galactic wind and the high velocity clouds in our Galaxy. The models of coronal formation which have been received most favourably are the galactic fountain models⁶³⁻⁶⁵ which describe the flow of hot gas from the plane into the halo where it cools radiatively, condenses forming clouds due to thermal instabilities, and falls back to the plane. The advantage of these models is that the high velocity clouds may be identified with the condensations and Bregman⁶⁴ has claimed that by assuming an isothermal corona with a radius dependant density proportional to the total density of neutral hydrogen, the velocity distribution of the high velocity and the intermediate velocity clouds can be understood. He envisages that near the plane the rising gas is accelerated radially by the coronal pressure gradient. Eventually the gas travels high enough above the plane that it has had time to cool to a temperature at which thermal instabilities occur and the clouds, which consequently form, fall like ballistic particles back to the plane after striking it near the point at which the hot gas from which they formed originally left. The

Table 3 Observed frequency of narrow quasar absorption lines

Ion	$\Delta z > 0.1$			$\Delta z < 0.1$		
	Actual detections	Total detections possible	Ratio (%)	Actual detections	Total detections possible	Ratio (%)
C ⁺	19	48	39.6	4	23	17.4
C ³⁺	62	63	98.4	23	24	95.8
N ⁴⁺	4	46	8.7	16	24	66.7
O ⁰	9	43	20.9	2	24	8.3
O ⁵⁺	2	5	40.0	2	3	66.7
Si ⁺	23	61	37.7	5	23	21.7
Si ²⁺	18	37	48.6	7	22	31.8
Si ³⁺	27	55	49.1	14	23	60.9

very major problems with Bregman's model are that the temperature at the base of the corona must lie between 7×10^5 K and 1.6×10^6 K, in conflict with the N V and O VI observations, and that it is very difficult to understand why 8×10^4 K halo gas would be very abundant in such a model as the cooling time is at its shortest near that temperature⁴⁰. Specifically, N V should be more abundant than Si IV if the fountain model were correct.

It is tempting to construct galactic coronal models similar to those used for solar and stellar coronae. However, a major difference between the solar and galactic coronae is that the structure of the solar corona is governed by thermal conductivity while at 8×10^4 K the galactic corona is too cool for conduction to be significant. Recently, Sturrock and Stern⁶⁶ have proposed that the galactic corona is heated by the dissipation of magnetic energy generated by galactic differential rotation. The generation of magnetic energy in the plane produces regions with high magnetic pressure which are buoyant and rise to the halo where field reconnection occurs dissipating the magnetic energy. This model can produce a 10^6 K corona but cannot produce 8×10^4 K gas. Book⁶⁷ has suggested that the work done by the expansion of such magnetic bubbles can heat the solar corona without the need for wave or current dissipation. In the case of the galactic corona the detailed nature of the dissipation of a magnetic bubble's energy will fix the ratio of the bubble formation time and the length of time required to dissipate its energy. If the ratio is nearly unity the corona will be approximately in steady-state.

It has been suggested^{48,68} that the dissipation in the halo gas of Alfvén waves produced in the disk by any source including the expansion and fragmentation of supernova remnants can heat the galactic halo gas. Dissipation can occur through ordinary viscosity or through nonlinear decay into a sound wave and another Alfvén wave. In many cases the latter mechanism may dominate. If it does the heating mechanism turns off rapidly when the temperature becomes too high, because conductive dissipation of sound waves will become rapid preventing them from growing and will lead to a great weakening of the nonlinear wave-wave interaction. For reasonable choices of parameters this rapid turnoff can occur near 8×10^4 K. Such structure in the heating rate probably explains why the gas is thermally stable even though the cooling rate coefficient is structureless between 6×10^4 K and 2×10^5 K (refs 68, 69). Hartquist⁴⁸ has shown that a low density, slow wind, driven by wave pressure will have a temperature which varies slowly with density. It is possible that the galactic corona is a wave-driven fountain. Thermal instability will occur at heights at which the wave energy is no longer large, resulting in cloud formation. As mentioned previously, the Al III data may require that the clouds form from gas with $T \leq 5 \times 10^4$ K, but the Al III also may originate in other cooling gas such as that produced by the collision of two initially cool clumps in high latitude clouds⁴⁵. The velocity dispersion of the clumps is so high that they are not virially stable, but they are probably contained by magnetic pressure. The large scale field strength in the corona required to achieve the containment is comparable with that necessary for the wave heating models to be viable⁴⁵.

Narrow line absorption systems in quasar spectra

The suggestion of Spitzer and Bahcall⁷⁰ that many of the narrow line absorption systems seen in the spectra of quasars may arise in the coronae of intervening galaxies has received observational support^{43,46,71-75}. If the narrow line systems do originate in the intervening galaxies, analysis of them can yield information about the evolution of the interstellar medium and the structure of galaxies. In addition the study of absorption features like those of O VI (λ 1,033 Å and λ 1,037 Å) arising in hot gas and with rest wavelengths which are currently inaccessible, may provide insight into the structure of the corona of our Galaxy.

To provide a further test of the intervening galaxy hypothesis, we have used the observed abundances in the hot galactic disk and halo gas to estimate the column densities of C³⁺, Si³⁺, N⁴⁺ and O⁵⁺ along sight lines through the Galaxy. We used these column densities when trying to understand the observed frequencies of detection of features in quasar absorption line systems. These frequencies of occurrence were determined from a study of all available data obtained with the University College London Image Photon Counting System (IPCS)⁷⁶.

To estimate the column densities observed along a line of sight perpendicular to the disk of our Galaxy we considered two separate contributions. Jenkins¹¹ found that the O⁵⁺ density distribution has a scale height of 300 pc. Hence, for each species we have included a disk contribution given by the average disk abundance multiplied by 600 pc. Observations reporting the average densities of C IV and Si IV in the halo along a number of sight lines were discussed earlier. The density of N V in the halo is uncertain, but excluding three lines of sight, its number density is $< 1.5 \times 10^{-9}$ cm⁻³. Instruments capable of detecting O VI in the halo have not been built. Table 2 summarizes the existing data on these four ions and estimates their total column densities. We have assumed that the halo gas has a scale height of 3 kpc. The densities in Table 2 are averages from several lines of sight. Interstellar material is fairly inhomogeneous. For example, the column densities of material obtained by Savage and de Boer^{39,40} towards the Small Magellanic Cloud are twice those they obtained towards the Large Magellanic Cloud. However, the column densities in Table 2 provide a reasonable data base to compare with QSO absorption data.

Column densities of quasar absorption lines are available for a few exceptionally well studied rich systems. These few column densities are in reasonable agreement with the column densities found in halo material towards the Magellanic cloud stars⁷⁵, but these comparisons were primarily for Si IV, C IV, and features of more lowly ionized species.

Rather than consider column densities, we will concentrate on the frequencies with which features occur. For each absorption line we compared the number of actual detections of a feature with the number of systems for which observations with adequate signal-to-noise have been made in the appropriate wavelength range. With very high quality data a comparison of the equivalent widths of various features seen in a system to that of Si IV would be possible. As such data do not exist,

we are fortunate that, as discussed below, the detection limits are low enough to be comparable with those needed to see Si IV, a moderately abundant ion in our Galaxy and in the absorption line systems. We expect ions less abundant than Si IV to be detected rarely and those much more abundant than Si IV to be detected nearly always. The IPCS data we used were for 30 different quasars^{72,76-90} and show 90 definite absorption systems. For 27 of these systems, the differences, Δz , between their redshift, z_a , and the quasar emission line redshift, z_e , is <0.1 . We treated this group separately as they may be associated with or be near to the quasar so that their ionization structure is influenced by the quasar radiation field. The results of this analysis are presented in Table 3.

The differences between the two groups of systems are striking. It has been noted previously that systems with small Δz often contain N V whereas other systems do not⁹¹. This effect is demonstrated in Table 3. In addition there are few small Δz systems in which C II is detected and the silicon ionization balance in those systems seems to favour more highly ionized species. These trends may be due to the effects of high photoionization rates. Some of the small Δz systems not showing these shifts may be near quasars but shielded by gas in which another absorption line system is formed.

The results in Table 2 can be used to understand the data for the group of absorption lines with $\Delta z > 0.1$. In making this comparison it should be remembered that the smallest detectable equivalent width does vary from quasar to quasar and is a function of wavelength for the systems studied. In addition, it is sometimes difficult to establish unique identifications for lines in the Ly α forest ($\lambda < (1+z_e)1,216 \text{ \AA}$). In particular, for $\Delta z > 0.1$ the N V lines lie in the Ly α forest. Knowing these difficulties, we can still draw a few general conclusions.

Si IV is detected in roughly one half of the $\Delta z > 0.1$ systems, implying that the typical Si³⁺ column density is roughly the minimal detectable one if most systems originate in relatively similar objects. Apparently, Si²⁺ which is also seen in ~50% of the systems has a column density comparable with that of Si³⁺ as is expected in a limited temperature range around 10^5 K . Thus the silicon absorption line data show that the temperature is perhaps marginally higher in most gas seen in absorption against quasars than in the hot galactic halo gas.

Table 2 shows that the N⁴⁺ column densities may be expected to be as much as five times lower than those of Si³⁺ if N⁴⁺ exists only in the disks of galaxies. In addition the oscillator strength for the observable transitions of N⁴⁺ is about 3 times weaker than that of the transitions for Si³⁺. However, in those few systems in which N V is seen its feature is of at least comparable strength to that of Si IV. Clearly, the N V detections cannot be understood if we assume that the column densities in Table 2 are correct for all intervening galaxies. Smith and Hartquist¹³ found considerable scatter in the disk N V number density, which can be caused by relatively small changes of temperature around $T = 10^{5.4} \text{ K}$. Pettini and West⁴² have also found some scatter in the halo number densities of C IV and Si IV (though little scatter in their ratio). It is probably possible to understand the very few detections in quasar absorption line systems by

assuming that N V exists only in the disks of intervening galaxies and that its density in those disks has a distribution not too different than that in the galactic disk. However, from these data we cannot rule out the presence of interstellar N V in the haloes of the intervening galaxies. Alternatively, the N V absorption features may arise in somewhat unusual regions; perhaps they are in galaxies in their early stages of evolution.

C³⁺ is observed in a high percentage of the absorption line systems. Its column densities in the galactic halo are about 4 times that of Si³⁺, but the relevant oscillator strength is about 3 times weaker than that of Si³⁺. For C IV features to be observed in so many systems the density of C³⁺ to Si³⁺ must in general be >4.5 implying as did the Si III to Si IV data that the temperature of the intervening gas is marginally greater than in our galactic halo. An observational bias favouring C IV detection exists: the presence of the C IV doublet is used often to establish the reality of an absorption system.

Data for O VI features are scarce. In our Galaxy the contribution to the O⁵⁺ column density from the disk gas is greater than the total column density of Si³⁺, but the relevant oscillator strengths are in the ratio of 1:4. Even if O⁵⁺ is present only in the disks of our Galaxy and of the intervening galaxies we would expect reasonably frequent detection of it. If the galaxies have densities of O⁵⁺ in their haloes comparable with that found in the galactic disk, one would expect nearly a 100% detection rate. We cannot draw strong conclusions from the small number of systems observed for O VI, but on the basis of the ~50% detection rate in the $\Delta z > 0.1$ systems, we suspect that our Galaxy has a very low abundance of O⁵⁺ in its halo interstellar gas.

Future work

The existing UV observations of the galactic corona have provided much information about the physics of the hot interstellar gas. The theoretical explanations of the existing data including the N V and O VI measurements in the disk seem incomplete, and more work is necessary to explain them and the soft X-ray emission. Additional observations of the hot halo gas, whose properties are not fully understood, and of the quasar absorption line regions are desirable. Searches for O VI features in the galactic halo gas will have to wait until the next generation of UV observatories flies, but further optical searches for N V and O VI in quasar absorption line systems are possible. Eventually, they and the future galactic UV O VI observations may contribute significantly to the determination of the origin of the quasar absorption line systems.

Despite the apparent similarity of the physical properties of the galactic coronal gas to the gas in quasar absorption line regions, the question of whether the lines of sight to quasars cross enough galaxies to explain the number of detected absorption systems remains^{82,91} as the hot galactic gas seems to be fairly close to the disk. It is encouraging that the physical properties of the coronal gas of the Large Magellanic Cloud appear to be so similar to the galactic coronal gas⁹² suggesting that galaxies with a very wide range of masses contribute to the absorption by intervening galaxies.

- Spitzer, L. Jr *Astrophys. J.* **124**, 20-34 (1956).
- Hartquist, T. W., Snijders, M. A. J. & West, K. A. *Mon. Not. R. astr. Soc.* (submitted).
- Baliunas, S. L. & Butler, S. E. *Astrophys. J. Lett.* **235**, L45-L48 (1980).
- Shapiro, P. R. & Moore, R. L. *Astrophys. J.* **207**, 460-483 (1976).
- Allen, C. W. *Astrophysical Quantities* 3rd edn (Athlone, London, 1973).
- Jenkins, E. B. & Meloy, D. A. *Astrophys. J. Lett.* **193**, L121-L125 (1974).
- York, D. G. *Astrophys. J. Lett.* **193**, L127-L131 (1974).
- Jenkins, E. B. *Astrophys. J.* **219**, 845-860 (1978).
- Castor, J., McCray, R. & Weaver, R. *Astrophys. J. Lett.* **200**, L107-L110 (1975).
- Weaver, R., McCray, R. & Castor, J. *Astrophys. J.* **218**, 377-395 (1977).
- Jenkins, E. B. *Astrophys. J.* **220**, 107-123 (1978).
- York, D. G. *Astrophys. J.* **213**, 43-51 (1977).
- Smith, L. J. & Hartquist, T. W. *Mon. Not. R. astr. Soc.* **192**, 73P-77P (1980).
- Bruhweiler, F. C., Kondo, Y. & McClusky, G. E. *Astrophys. J.* **237**, 19-25 (1980).
- Black, J. H., Dupree, A. K., Hartmann, L. W. & Raymond, J. C. *Astrophys. J.* **239**, 502-514 (1980).
- Cowie, L. L., Taylor, W. & York, D. G. *Astrophys. J.* **248**, 528-541 (1981).
- Spitzer, L. Jr *Physical Processes in the Interstellar Medium* (Wiley, New York, 1978).
- Chandrasekhar, S. & Fermi, E. *Astrophys. J.* **118**, 113-115 (1953).
- Krall, N. A. & Trivelpiece, A. W. *Principles of Plasma Physics* (McGraw-Hill, New York, 1973).
- Rosenbluth, M. N. & Longmire, C. L. *Ann Phys.* **1**, 120-140 (1957).
- Draine, B. T. *Astrophys. J.* **241**, 1021-1038 (1980).
- Smith, L. J., Willis, A. J. & Wilson, R. *Mon. Not. R. astr. Soc.* **191**, 339-347 (1980).
- Laurent, C., Paul, J. A. & Pettini, M. *Astrophys. J.* **260**, 163-182 (1982).
- Williamson, F. O. et al. *Astrophys. J. Lett.* **193**, L133-L137 (1974).
- Fried, P. M., Nousek, J. A., Sanders, W. T. & Kraushaar, W. L. *Astrophys. J.* **242**, 987-1004 (1980).
- Rosner, R. et al. *Astrophys. J. Lett.* **249**, L5-L10 (1981).
- Levine, A., Rappaport, S., Doxsey, R. & Jernigan, G. *Astrophys. J.* **205**, 226-232 (1976).
- Cash, W., Maline, R. & Stern, R. *Astrophys. J. Lett.* **204**, L7-L11 (1976).
- Yentis, D. J., Novick, R. & Vander Bout, P. *Astrophys. J.* **177**, 365-386 (1972).
- Stern, R. & Bowyer, S. *Astrophys. J.* **230**, 755-767 (1979).
- Paresce, F., McKee, C. F. & Bowyer, S. *Astrophys. J.* **240**, 387-400 (1980).
- Jura, M. *Astrophys. J.* **227**, 798-800 (1979).
- Code, A. D., Davis, J., Bless, R. C. & Hanbury Brown, R. *Astrophys. J.* **203**, 417-434 (1976).
- Nandy, K., Thompson, G. I., Jamar, C., Monfils, A. & Wilson, R. *Astr. Astrophys.* **44**, 195-203 (1975).
- Nandy, K., Thompson, G. I., Jamar, C., Monfils, A. & Wilson, R. *Astr. Astrophys.* **51**, 63-69 (1976).
- Seaton, M. J. *Mon. Not. R. astr. Soc.* **187**, 73P-76P (1978).
- Overbeck, J. W. *Astrophys. J.* **141**, 864-866 (1965).
- Hayakawa, S. *Prog. Theor. Phys.* **43**, 1224-1230 (1970).

39. Savage, B. D. & de Boer, K. S. *Astrophys. J. Lett.* **230**, L77–L82 (1979).
40. Savage, B. D. & de Boer, K. S. *Astrophys. J.* **243**, 460–484 (1981).
41. Bromage, G. E., Gabriel, A. H. & Sciamia, D. W. in *2nd European IUE Conf. ESA-SP 157*, 345–351 (1980).
42. Pettini, M. & West, K. A. *Astrophys. J.* (in the press).
43. Ulrich, M.-H. *et al. Mon. Not. R. astr. Soc.* **192**, 561–580 (1980).
44. Savage, B. D. & Fitzpatrick, E. L. *Bull. Am. astr. Soc.* **13**, 895 (1981).
45. Dyson, J. E. & Hartquist, T. W. *Mon. Not. R. astr. Soc.* (submitted).
46. York, D. G. *Ann. Rev. Astr. Astrophys.* **20**, (in the press).
47. Pettini, M., Hartquist, T. W. & Tallant, A. *Mon. Not. R. astr. Soc.* (submitted).
48. Hartquist, T. W. *Mon. Not. R. astr. Soc.* (in the press).
49. Verschuur, G. L. *Ann. Rev. Astr. Astrophys.* **13**, 257–293 (1975).
50. Bone, D., Hartquist, T. W. & Sanford, P. *Astrophys. Spac. Sci.* (in the press).
51. Dickey, J. M., Salpeter, E. E. & Terzian, Y. *Astrophys. J. Lett.* **211**, L77–L81 (1977).
52. Hartquist, T. W., Black, J. H. & Dalgarno, A. *Astrophys. J.* **259**, 591–594 (1982).
53. Botcher, C., McCray, R. A., Jura, M. & Dalgarno, A. *Astrophys. Lett.* **6**, 237–241 (1970).
54. Cox, D. P. & Smith, B. W. *Astrophys. J. Lett.* **189**, L105–L108 (1974).
55. McKee, C. F. & Ostriker, J. P. *Astrophys. J.* **218**, 148–169 (1977).
56. Weisheit, J. C. & Collins, L. A. *Astrophys. J.* **210**, 299–310 (1976).
57. Sciamia, D. *Nature* **240**, 456–457 (1972).
58. Sagdeev, R. Z. & Galeev, A. A. *Nonlinear Plasma Theory* (Benjamin, New York, 1969).
59. Wentzel, D. G. *Astrophys. Lett.* **10**, 167–170 (1972).
60. Wentzel, D. G. A. *Rev. Astr. Astrophys.* **12**, 71–96 (1978).
61. Chevalier, R. A. & Gardner, J. *Astrophys. J.* **192**, 457–463 (1974).
62. Chevalier, R. A. & Oegerle, W. R. *Astrophys. J.* **227**, 398–406 (1979).
63. Shapiro, P. R. & Field, G. B. *Astrophys. J.* **205**, 762–765 (1976).
64. Bregman, J. N. *Astrophys. J.* **236**, 577–591 (1980).
65. Kahn, F. in *Investigating the Universe* (ed. Kahn, F.) 1–28 (Reidel, Dordrecht, 1981).
66. Sturrock, P. A. & Stern, R. *Astrophys. J.* **238**, 98–102 (1980).
67. Book, D. L. *Comments Plasma Phys. Controlled Fusion* **6**, 193–198 (1981).
68. Hartquist, T. W. & Tallant, A. *Mon. Not. R. astr. Soc.* **196**, 527–532 (1981).
69. Summers, H. P. & McWhirter, R. W. P. *J. Phys.* **B12**, 2387–2412 (1979).
70. Bahcall, J. N. & Spitzer, L. Jr *Astrophys. J. Lett.* **156**, L63–L65 (1969).
71. Boksenberg, A. *Phys. Scr.* **17**, 205–214 (1978).
72. Young, P. J. *et al. Astrophys. J.* **229**, 891–908 (1979).
73. Sargent, W. L. W., Young, P. J., Boksenberg, A., Carswell, R. F. & Whelan, J. A. J. *Astrophys. J.* **230**, 49–67 (1979).
74. Snijders, M. A. J. *ESA SP-157, 2nd European IUE Conf.* lxxi–lxxx (1980).
75. Savage, B. D. & Jeske, N. A. *Astrophys. J.* **244**, 768–776 (1981).
76. Boksenberg, A. & Burgess, D. E. *Proc. Symp. on Astronomical Observations with Television Type Sensors*, British Columbia (eds Glaspey, J. W. & Walker, G. H. A.) 21–30 (1973).
77. Boksenberg, A., Carswell, R. F., Smith, M. G. & Whelan, J. A. J. *Mon. Not. R. astr. Soc.* **184**, 773–782 (1978).
78. Wright, A. E., Morton, D. C., Peterson, B. A. & Jauncey, D. L. *Mon. Not. R. astr. Soc.* **189**, 611–620 (1979).
79. Blades, J. C., Murdoch, H. S. & Hunstead, R. W. *Mon. Not. R. astr. Soc.* **191**, 61–68 (1980).
80. Blades, J. C., Hunstead, R. W. & Murdoch, H. S. *Mon. Not. R. astr. Soc.* **194**, 669–678 (1981).
81. Morton, D. C., Chen, J. S., Wright, A. E., Peterson, B. A. & Jauncey, D. L. *Mon. Not. R. astr. Soc.* **193**, 399–414 (1980).
82. Sargent, W. L. W., Young, P. A., Boksenberg, A. & Tytler, D. *Astrophys. J. Suppl.* **42**, 41–81 (1980).
83. Chen, J. S., Morton, D. C., Peterson, B. A., Wright, A. E. & Jauncey, D. L. *Mon. Not. R. astr. Soc.* **196**, 715–730 (1981).
84. Blades, J. C., Hunstead, R. W., Murdoch, H. S. & Pettini, M. *Mon. Not. R. astr. Soc.* (in the press).
85. Carswell, R. F., Whelan, J. A. J., Smith, M. G., Boksenberg, A. & Tytler, D. *Mon. Not. R. astr. Soc.* **198**, 91–110 (1982).
86. Pettini, M. *et al. Preprint Astrophys. J.* (submitted).
87. Sargent, W. L. W., Young, P. & Boksenberg, A. *Astrophys. J.* **252**, 54–68 (1982).
88. Wright, A. E., Morton, D. C., Peterson, B. A. & Jauncey, D. L. *Mon. Not. R. astr. Soc.* **199**, 81–91 (1982).
89. Young, P., Sargent, W. L. W. & Boksenberg, A. *Astrophys. J.* **252**, 10–31 (1982a).
90. Young, P., Sargent, W. L. W. & Boksenberg, A. *Astrophys. J. Suppl.* **48**, 455–505 (1982b).
91. Weymann, R. J., Carswell, R. F. & Smith, M. G. A. *Rev. Astr. Astrophys.* **19**, 41–76 (1981).
92. de Boer, K. S. & Savage, B. D. *Astrophys. J.* **238**, 86–92 (1980).
93. Levine, A., Rappaport, S., Halpern, J. & Walter, F. *Astrophys. J.* **211**, 215–222 (1977).

ARTICLES

MERLIN observations of superluminal radio sources

I. W. A. Browne, R. R. Clark*, P. K. Moore, T. W. B. Muxlow, P. N. Wilkinson, M. H. Cohen† & R. W. Porcas‡

Nuffield Radio Astronomy Laboratories, Jodrell Bank, Macclesfield, Cheshire SK11 9DL

† Owens Valley Radio Observatory, California Institute of Technology, Pasadena, California 91125, USA

‡ Max Planck Institut für Radioastronomie, Auf dem Hugel 69, D5300 Bonn 1, FRG

Powerful extragalactic radio sources may all be members of a single family in which those showing superluminal behaviour are the ones in which a relativistic jet is directed towards the observer. The superluminal sources are shown to have extended radio structures the form of which is consistent with such a unified scheme. This interpretation requires that the Hubble constant H_0 should be nearer to 100 than to 50 km s⁻¹ Mpc⁻¹.

OBSERVATIONAL evidence for superluminal velocities has so far been found in a small proportion of extragalactic radio sources¹. Only one of these superluminal sources has the steep radio spectrum² and symmetrical outer lobes characteristic of most 'normal doubles'. The emission from the other sources is dominated by flat radio spectrum compact cores. Little is known about the extended radio structure of such sources. We suggest nevertheless, that these flat spectrum core-dominated sources and the steep spectrum normal double sources are closely related. We show that there is extended structure surrounding the cores of most superluminal sources and that its properties are consistent with a unified scheme³ in which the core-dominated type sources happen to be those in which a relativistic jet is directed towards the observer.

We now present hybrid maps made with the Jodrell Bank Multi Element Radio Linked Interferometer Network (MERLIN) of nine superluminal sources and also the core-dominated quasar 4C39.25. The comparatively low observing frequencies of 408 MHz and 1,666 MHz are very suitable for showing up the steep spectrum emission typical of extended parts of radio sources. Maps made, using the techniques described in ref. 4, at 1,666 MHz have a resolution ~ 0.3 arc s and a dynamic range $\sim 1,000:1$, while those at 408 MHz have a resolution ~ 1 arc s and a dynamic range $\sim 200:1$.

The observational results on each source are summarized in Table 1. We take Hubble's constant $H_0 = 100$ km s⁻¹ Mpc⁻¹ and $q_0 = 0.05$. The factor h allows the results to be scaled for other values of H_0 . The MERLIN maps are displayed in Fig. 1.

Individual sources

NRAO 140: Marscher and Broderick⁵ have three epoch very-long baseline interferometry (VLBI) observations which show this source to be expanding at an angular rate between 0.10 and 0.14 m arc s yr⁻¹. The MERLIN 408-MHz map shows a weak (0.14 Jy) extended component in position angle (PA) 149°, 7.6 arc s from the 3.57 Jy unresolved core. The source is variable at 408 MHz with mean flux density over the period 1975–79 of 3.12 Jy (ref. 6). There is thus no evidence for extended structure in addition to that visible in our map.

3C120: The latest results on the motion in the core of this Seyfert Type I galaxy are presented by Walter *et al.*⁷. The MERLIN 1,666-MHz map shows a jet with an extended knot 4 arc s from the core. On the 408-MHz map this jet continues a further 10 arc s, bending increasingly to the north as it does so. This jet is visible on the Very Large Array (VLA) maps of Balick *et al.*⁸ and is also just visible on our 1,666-MHz map. The 408-MHz map contains 3.9 Jy compared with a total flux density of ~ 6.5 Jy (ref. 6). This implies that there is a lot of missing low brightness structure. The VLBI position angle is -108° , that of the 4 arc s knot -99° , and that of the jet -75° .

* Present address: Electrical Engineering Department, University of New Hampshire, Durham, New Hampshire 03824, USA.

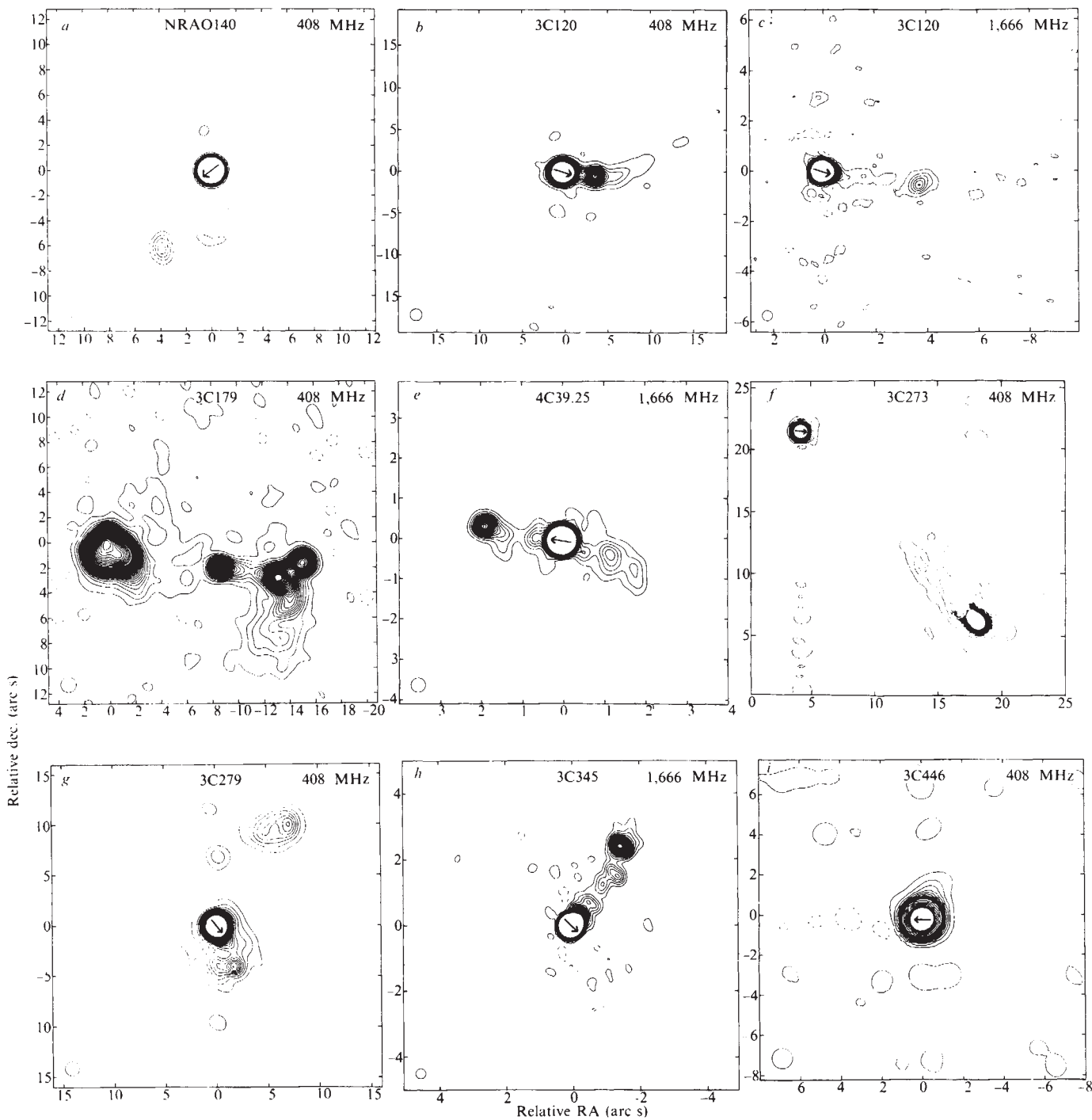


Fig. 1 MERLIN maps of superluminal sources with arrows showing the position angle of the smallest scale VLBI structure. *a*, NRAO 140 at 408 MHz (epoch 1981.6). Contours 0.5, 1.0, 1.5, 2.0, ... % of the peak brightness of 3.36 Jy per 1.0×1.0 arc s beam. *b*, 3C120 at 408 MHz (epoch 1981.4). Contours 0.25, 0.75, 1.25, 1.75, ... % of the brightness of 2.87 Jy per 1.4×1.4 arc s beam. *c*, 3C120 at 1,666 MHz (epoch 1981.0). Contours 0.1, 0.3, 0.5, 0.7, ... % of the peak brightness of 3.18 Jy per 0.4×0.4 arc s beam. *d*, 3C179 at 408 MHz (epoch 1981.5). Contours 1, 3, 5, 7, ... %, then 55, 60, 65, ... % of the peak brightness of 0.54 Jy per 1.2×1.2 arc s beam. *e*, 4C39.25 at 1,666 MHz (epoch 1980.7). Contours 0.33, 0.67, 1.00, 1.33, ... % of the peak brightness of 2.32 Jy per 0.35×0.35 arc s beam. *f*, 3C273 at 408 MHz (epoch 1981.5). Contours 0.1, 0.3, 0.5, 0.7, ... % of the peak brightness of 12.5 Jy per 0.9×0.9 arc s beam. *g*, 3C279 at 408 MHz (epoch 1981.4). Contours 0.5, 1.5, 2.5, 3.5, ... % of the peak brightness of 6.99 Jy per 1.4×1.4 arc s beam. *h*, 3C345 at 1,666 MHz (epoch 1980.9). Contours 0.1, 0.2, 0.3, 0.4, ... % of the peak brightness of 6.26 Jy per 0.25×0.25 arc s beam. *i*, 3C446 at 408 MHz (epoch 1981.7). Contours 0.5, 1.5, 2.5, 3.5, ... %, then 15, 20, 25% of the peak brightness of 9.50 Jy per 1.0×1.0 arc s beam. *j*, 3C454.3 at 1,666 MHz (epoch 1980.7). Contours 0.05, 0.10, 0.15, 0.20, ... % of the peak brightness of 8.46 Jy per 0.25×0.25 arc s beam. The knot at the end of the jet has a peak brightness of 0.20 Jy.

Table 1 Source properties

Name	Identification/redshift	θ (m arc s yr ⁻¹)	v/c	VLBI PA (Refs)	MERLIN PA	Magnetic field direction
NRAO 140	Q/1.258	0.13	5.4 h^{-1}	127 (5)	149	138
3C120	G/0.033	1.35	2.1 h^{-1}	-108 (7)	-99 to -75	-96
3C179	Q/0.846	0.14	4.2 h^{-1}	-88 (2)	-100	-96
4C39.25	Q/0.699	<0.02	<0.5 h^{-1}	81 (44)	80	?
3C273	Q/0.158	0.76	5.3 h^{-1}	-99 to -138 (45)	-138	-120
3C279	Q/0.538	0.5 or 1.5	(10 or 25) h^{-1}	-142 (7)	-157	-158
3C345	Q/0.595	0.36	8.2 h^{-1}	-135 (18) to -57 (45)	-38 to -32	-47
BL Lac	B/0.07	?	?	-170 (22)	—	133 or -47
3C446	Q/0.847	?	?	-90 (24)	-32	-84
3C454.3	Q/0.859	?	?	-65 (28)	-48	-64

The values of θ and v/c are taken from ref. 1 and the magnetic field direction from the intrinsic electric field position angles given by Simard-Normandin *et al.*⁴².

Note that the sense of jet curvature remains the same on all angular size scales.

Direct images of the galaxy have been obtained by Arp⁹ and also by Wlérick *et al.*¹⁰ who describe an optical continuum jet which emerges from the nucleus roughly parallel with the radio jet and which bends north at about the position of the 4 arc-s knot. The bend is in the same sense as the radio but is significantly larger.

Baldwin *et al.*¹¹ have mapped the velocity field of the galaxy and find an axis of symmetry nearly parallel to the VLBI axis. One possible interpretation of their observations is that the emission line gas is rotating about this axis. This is, however, difficult to reconcile with the detection of superluminal motion many explanations of which require the VLBI axis to be close to the line of sight. The true rotation velocities which would be implied if this were the case (≥ 600 km s⁻¹ for angles to the line of sight $< 30^\circ$) seem unrealistically large for rotating gas in a galaxy. We suggest that the gas is probably not rotating about the VLBI axis. The alternative explanations put forward by Baldwin *et al.* of either infalling or outflowing gas seem more plausible.

3C179: Porcas² finds the central component of this source to be expanding with $v = 4.2c$ h^{-1} . The MERLIN 408-MHz map shows strong double lobes and a jet pointing to the western lobe. **3C273:** MERLIN 408-MHz observations have been discussed by Conway *et al.*¹². The map presented here is a new one made with all six telescopes.

3C279: This is the first source in which rapid changes in structure were observed¹³ but the details of the structural changes are still not clear. Cohen and Unwin¹ suggest an expansion speed of $10c$ h^{-1} , while Pauliny-Toth *et al.*¹⁴ think the expansion rate might be as high as $25c$ h^{-1} . The MERLIN 408-MHz map shows the strong core and two extended outer components. On a higher resolution map the extension from the nucleus in PA -145° is clearly seen as a jet (Perley, personal communication). Our map contains virtually all the flux density. The weak component lying ~ 5 arc s due north of the core may be a sidelobe.

3C345: Recent VLBI results are reported in refs 15–17. The MERLIN 1,666-MHz map shows a curved jet starting from the nucleus and ending in a compact knot. Figure 2 shows the position angle of source elongation plotted as a function of

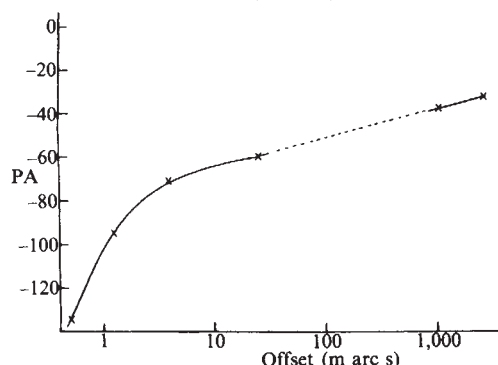


Fig. 2 The bend in 3C345. The position angle of a point in the jet as seen from the core is plotted against the distance of that point from the core. VLBI data are taken from refs 16, 18 and 45.

distance from the core. There are two things to note; first most of the bending occurs in the first 0.1% of the length of the jet well within the region in which the superluminal motion is observed. Second, the sense of curvature on the m arc s scale is retained on the arc s scale. This latter observation suggests that the cause of bending is the same on all scales. Browne *et al.*²⁰ have published a MERLIN 408-MHz map and they point out that 3C345 must have some low brightness emission on a scale of ≥ 5 arc s in addition to the jet.

BL Lac: Mutel *et al.*²¹ and Phillips and Mutel²² report superluminal behaviour in this source. We have no evidence for the source being resolved with MERLIN; at 408 MHz the flux density of any extended structure is $< 2\%$ of that of the core and at 1,666 MHz it is $< 0.5\%$ of the core flux density.

3C446: The MERLIN 408 MHz map shows a strong, unresolved component with a slight extension in PA -32° , in agreement with the 327-MHz lunar occultation results of Joshi and Gopal-Krishna²³. At higher frequencies VLA results²⁴ indicate a resolved component at PA 112° and separation ~ 0.6 arc s from the core. This is more closely aligned with the VLBI structure which is in PA 90° (ref. 24).

3C454.3: No direct structural changes have been observed in this source but rapid low frequency variability provides strong evidence for superluminal behaviour^{25,26}. To first order the VLBI structure consists of an unresolved core with a bright knot 7 m arc s distant in PA -65° (refs 27, 28). There is also some evidence for a weak VLBI component on the opposite side of the core. The MERLIN 1,666-MHz map shows a very weak, slightly curved jet leading to relatively bright component ~ 4 arc s away in PA -48° . The curvature along the MERLIN jet has the same sense as that between the jet and the VLBI nucleus.

4C39.25 (0923+392): This is a quasar whose VLBI structure has been monitored for several years and a limit on the expansion speed $\leq 0.5c$ h^{-1} can be set¹. The MERLIN 1,666-MHz map shows a dominant unresolved core and other complicated extended structure fairly symmetrically disposed about the core. The overall morphology strongly resembles that of 1150+597 (ref. 29) which also has a compact hot spot on one side of a strong core and a more relaxed structure on the other side; both sources also possess a steep spectrum halo^{29,20}. Higher resolution observations of 1150+497 clearly reveal a jet linking the core to the hotspot. Our observations are suggestive of a corresponding jet in 4C39.25 although a higher resolution map will be required to confirm this.

Discussion

Many different models have been proposed to explain the apparent superluminal motion in radio sources. Marscher and Scott³⁰ review most of the theories and find that few are free from conflict with observation. Of the ones that do survive most observational tests, those in which there is bulk relativistic motion of the radiating material are the most widely accepted. We will concentrate on this type of model.

Source morphology

The most obvious thing apparent from our maps is that every source except BL Lac has large scale structure in addition to its compact core. There is clearly great diversity in the struc-

tures; some of them have one-sided jets (3C120, 3C179, 3C273, 3C345 and 3C454.3) and some have extended structure on both sides of the core (3C179, 3C279 and 4C39.25). At least with the present dynamic range, NRAO 140 appears to have a single isolated region of emission well away from the unresolved core. The flat spectrum sources amongst the superluminals have structures very much like other sources with similar spectra^{20,31}, and the only steep spectrum source 3C179 has a fairly normal double structure. Superluminal sources therefore do not look like some completely separate class of exotic object.

Asymmetries

One-sided jets have been observed in extragalactic sources having a wide range of luminosities³². There is considerable evidence that the jets in low luminosity radio galaxies are non-relativistic³³ but the situation is not so clear for the more powerful sources³⁴.

The MERLIN maps of the superluminal sources (except BL Lac) all show a certain degree of asymmetry about the central core component, ranging from a weak jet connecting the core and the western lobe in 3C179 to a totally one sided jet in 3C273. Six of the sources have clearly defined asymmetric VLBI structure (NRAO 140, 3C120, 3C179, 3C273, 3C345 and 3C454.3) and in all these cases the asymmetry in the extended structure is in the same sense. This suggests that the mechanism responsible for the one-sidedness in the VLBI cores is also responsible for that on the arc second scale²⁹. Clearly in the cores there is strong evidence for relativistic velocities which could easily account for the observed one-sidedness even if the sources were intrinsically two-sided. Do the relativistic velocities continue into the arc second jets? Evidence that they do comes from the results discussed by Browne *et al.*²⁰ and P. K. Moore (in preparation) concerning the absence of isolated large-scale jets (that is, those not associated with a strong core). The lack of such jets implies that, if the core emission is Doppler beamed, the jet emission has to be also and so this points towards relativistic jet velocities which continue from the cores to the outer parts of the sources. Doppler beaming is sufficient to explain the observed asymmetries, although relativistic jets which are intrinsically one-sided cannot be ruled out.

Unified scheme

As well as one-sided, sometimes jet-like features, four of the sources have still more extended structure. In 3C179 and 3C279 this takes the form of well defined lobes, but in 3C120 and 3C345 its presence is only betrayed by maximum interferometer visibilities significantly less than unity, even on the shortest baselines. The existence of the extended emission is particularly significant as it is most unlikely to be highly beamed, whereas the core emission might well be. If this is so there should exist a much larger population of unaligned objects dominated by their steep spectrum unbeamed emission. Such considerations led to the proposal of a unified scheme by Orr and Browne³ in which these unaligned counterparts of superluminal sources (and other core-dominated sources) are hypothesized to be normal steep spectrum double sources. Such an idea was dismissed by Scheuer and Readhead³⁵ because they deduced that the Lorentz factor (γ) of the core material was too low ($\gamma < 2$). In fact, Orr and Browne show that the statistical properties of quasars are compatible with a mean core Lorentz factor ~ 5 and therefore such a scheme is viable.

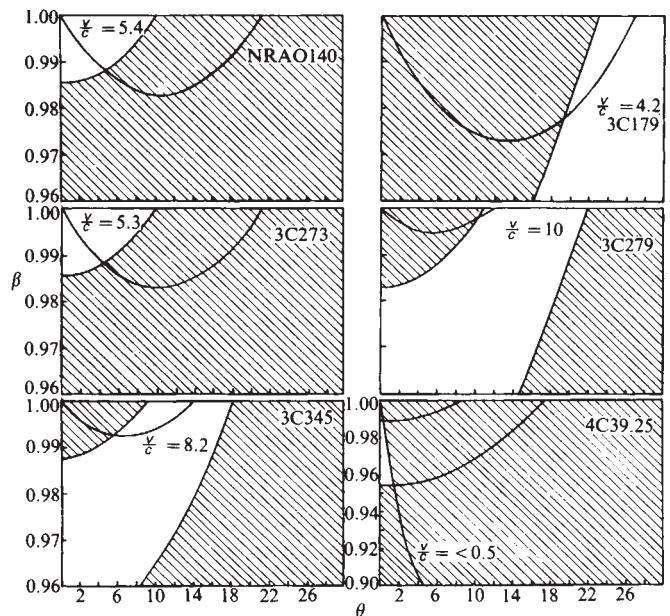


Fig. 3 θ/β diagrams for six sources with measured expansion velocities or limits on expansion velocities. The unified scheme constrains the sources to lie within the unshaded portions of the diagrams. The observed expansion rate means that the source must also lie on the v/c line. The following values of $R(\theta)$ have been used: > 100 for NRAO 140, 0.5 for 3C179, > 100 for 3C273, 10 for 3C279, 20 for 3C345 and 23 for 4C39.25.

For the sources in the present sample we can make quantitative tests to see if their properties are consistent with the unified scheme. Orr and Browne assume that the intrinsic ratio between the strength of the core and the lobes, defined as

$$R_T = \frac{\text{Core flux density when seen with its axis perpendicular to the line of sight}}{\text{The lobe flux density}}$$

is nearly a constant. The observed ratio $R(\theta)$, which has a very large spread from source to source will, of course, depend on the aspect of the source.

Following van Groningen *et al.*³⁶ we have plotted θ/β diagrams (Fig. 3) for those five quasars with well defined rates of expansion. θ is the angle which the motion in the core makes with the line of sight and βc is the velocity of the radiating component. In relativistic beam models an observed superluminal motion with velocity v requires the material to be moving with speed βc given

$$\beta = \frac{v/c}{\sin \theta + v/c \cos \theta}$$

Each object is constrained by the observed rate of expansion to lie on a line in the θ/β plane given by the above equation. (We assume $h = 1$ for these diagrams; see above.)

If we assume a specific value for R_T we can plot another line in the θ/β plane on which the quasar should lie using the equation³ $R(\theta) = \frac{1}{2} R_T (1 - \beta \cos \theta)^{-2} + \frac{1}{2} R_T (1 + \beta \cos \theta)^{-2}$. However, to allow for some realistic dispersion in R_T we plot two lines corresponding to $R_T = 0.006$ and $R_T = 0.10$. These refer to a frequency of 5 GHz and are respectively a factor of 4 lower and a factor of 4 higher than Orr and Browne's³ preferred value of 0.024. (The range $0.006 < R_T < 0.1$ is small compared with the observed spread in $R(\theta)$.) As can be seen from Fig. 3 in which the line defined by the expansion rate crosses an unshaded area, most of the sources are consistent with the scheme, although in the cases of 3C273 and NRAO 140 this is only by virtue of upper limits to the strength of any lobe-like structure. Should new observations show the lobe strengths very much weaker than the present limits the scheme, at least in its present simple form, would fail. We also note that 3C279 has relatively strong lobes compared with the core and

Table 2 Deprojected linear sizes for six quasars

Source	Observed angular size (arc s)	Allowed angle to the line of sight (deg)	Deprojected linear sizes (kpc)
NRAO 140	7.5	> 5	> 482
3C179	15.5	19–29	235–160
3C273	21	< 4.6	> 482
3C279	15	10–12	351–297
3C345	2.8	5.5–14	126–52
4C39.25	4	< 1.4	> 760

only just fits for the lower of the two possible values of v/c (that is, 10). If $v/c = 25$, as has been suggested by Pauliny-Toth *et al.*¹⁴ then it does not fit.

We can use the θ/β plots in Fig. 3 to define allowable ranges of θ for each source and hence work out limits on their deprojected sizes. We assume that the same value of θ applies to the large scale source structure as to the cores. Such an assumption is probably justified for large values of θ ($\geq 10^\circ$) as the work of

Macklin³⁷ on classical doubles indicates that the average bends in these are $\leq 5^\circ$, but for the sources with small θ our estimates are likely to have large errors. Table 2 lists these deprojected linear size estimates. All the values are within the range known for normal double quasars (≤ 550 kpc (refs 38, 39)) and so angular sizes present no problems for the scheme.

The fact that 3C179 is consistent with the unified scheme is interesting. If in 3C179 the core Lorentz factor is < 10 ($\beta < 0.995$) the scheme requires the angle to the line of sight, θ , to be in the range $19^\circ < \theta < 24^\circ$ giving a deprojected linear size between 167 and 136 kpc. This source illustrates that for not unrealistic Lorentz factors, superluminal motion can be observed in the cores of sources whose angles to the line of sight are quite large. If the unified scheme is correct, many cores of classical doubles should exhibit superluminal motion. That such motion has only been observed in one source is probably a reflection of the practical difficulty of making VLBI observations of weak cores.

Figure 3 also shows a θ/β plot for 4C39.25 using the limit on expansion velocity ($v/c < 0.5 h^{-1}$) given by Cohen and Unwin¹. This source would at first sight appear to be at odds with the scheme but for sufficiently small values of θ such a low apparent expansion speed will be observed even for relatively high values of γ . The probability of finding such a source is low but not prohibitively so; for $\gamma = 5$ the probability that a source selected by its core flux density having an angle to the line of sight $< 0.5^\circ$ is 0.4% (ref. 40). One problem with this interpretation is that the extended structure (Fig. 1) does not resemble a normal double source viewed almost directly along its axis; it appears to be too well collimated. This problem could be alleviated if there is some bending of the source axis which increases the angle to the line of sight of the outer structure. Such an effect would also reduce the deprojected linear size from the 2 Mpc it would be if the angle were truly 0.5° . We conclude that the above explanation for the lack of observed motion in 4C39.25 is not entirely out of the question.

Hubble's constant

Previously we used $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ($h = 1$) to plot Fig. 3. If we had used $h = 0.5$ instead this would have increased the values of the expansion rates by a factor of two and hence the required Lorentz factors. As far as the unified scheme is concerned this would make it more difficult to accommodate sources such as 3C279, which has strong extended structure, while making other sources like NRAO 140, 3C273 and 3C446 easier to explain. Individual objects therefore do not put very useful constraints on H_0 . However, a distance-independent average value of γ can be determined using the unified scheme by requiring it to account for the relative numbers of flat and steep spectrum objects. This value, which is well constrained by the observed properties of quasars, should be consistent with the average γ required to explain the observed superluminal motions. The unified scheme requires $\bar{\gamma} \sim 5$ which is only just consistent with the properties of the five known superluminal quasars if $h = 1$ (giving $\bar{\gamma}_{\text{observed}} \geq 7$). Much smaller values of h (such as $h = 0.5$ (ref. 41)) are excluded by the scheme. Conversely, if h is established to be as low as 0.5 the scheme fails.

Component spectra and correlations

Figure 4 shows spectral decompositions for the nine superluminal sources and 4C39.25. From 408 MHz to 50 GHz all the core spectra are relatively flat; the mean core spectral index is 0.01 ± 0.14 . The average spectral index of the extended structure is -0.8 typical of such structure in other powerful radio sources.

We have looked for correlations between the various source properties some of which are listed in Table 1. There is a good correlation between magnetic field direction, inferred from linear polarization measurements⁴² and the position angle of elongation of the whole source. Such an effect has been pointed out before in 3C273, 3C345 and 3C454.3 by Davis *et al.*⁴³ We

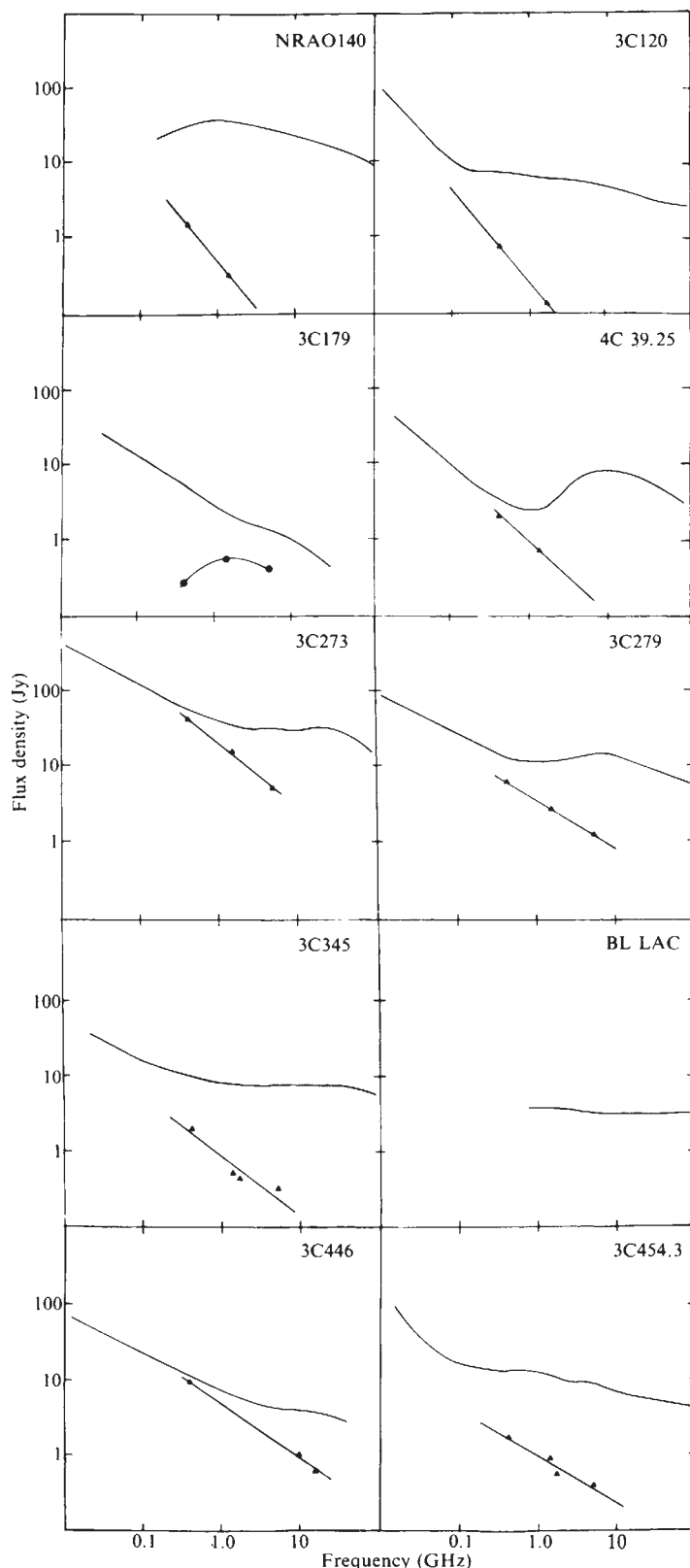


Fig. 4 Radio spectra of the superluminal sources. ▲, Flux density of extended emission. ●, Core flux density.

find that the magnetic field direction is often not perfectly aligned with axis of the dominant VLBI core, but usually has some position angle between that of the smallest and the largest scale structure. This suggests that most of the polarized flux density comes from the VLBI jet, not the core, and that the field direction follows the line of this jet as it does on the arc second scale²⁹.

Conclusions

Superluminal sources nearly all have structures on the scale of a few arc seconds which, although very different from source to source, share common morphological characteristics such as jets and double lobes with other powerful extragalactic objects. Such properties encourage the idea that all such sources are closely related and indeed the properties of most of the superluminal sources are consistent with Orr and Browne's unified scheme³. A consequence of such a scheme is that superluminal

motion should be detectable in most core-dominated sources and many steep spectrum relatively normal doubles like 3C179. A small percentage of core-dominated sources will not show rapid expansion because their angles to the line of sight are very small. 4C39.25 may be one of these.

To make further progress, systematic searches for superluminal motion in a complete sample of sources are essential. The unified scheme which makes well defined predictions for the distribution of apparent expansion velocities would stand or fall by such results.

We thank Tony Foley and Richard Davis for providing a new map of 3C273 and K. J. Johnston, R. A. Perley and G. Wlérick for providing data before publication. P.N.W. acknowledges the receipt of an SERC Advanced Fellowship. M.H.C. is grateful to the Guggenheim Memorial Foundation for a fellowship and to the Institute of Astronomy, Cambridge for their hospitality.

Received 15 June; accepted 11 August 1982.

1. Cohen, M. H. & Unwin, S. C. *IAU Symp. No. 97*, 345–354 (1982).
2. Porcas, R. W. *Nature* **294**, 47–49 (1981).
3. Orr, M. J. L. & Browne, I. W. A. *Mon. Not. R. astr. Soc.* **200**, 1067–1081 (1982).
4. Cornwell, T. J. & Wilkinson, P. N. *Mon. Not. R. astr. Soc.* **196**, 1067–1086 (1981).
5. Marscher, A. P. & Broderick, J. J. *IAU Symp. No. 97*, 359–360 (1982).
6. Fanti, C. et al. *Astr. Astrophys. Suppl.* **45**, 61–78 (1981).
7. Walker, R. C. et al. *Astrophys. J.* **257**, 56–62 (1982).
8. Balick, B., Heckman, T. M. & Crane, P. C. *Astrophys. J.* **254**, 483–488 (1982).
9. Arp, H. C. *Optical Jets in Galaxies* 53–61 (ESO/ESA Workshop 1981).
10. Wlérick, G., Bouchet, P., Cayatte, V. & Michet, D. *Astr. Astrophys.* **102**, L17–L20 (1981).
11. Baldwin, J. A. et al. *Astrophys. J.* **236**, 388–405 (1980).
12. Conway, R. G., Davis, R. J., Foley, A. R. & Ray, T. P. *Nature* **294**, 540–542 (1981).
13. Whitney, A. R. et al. *Science* **173**, 225–230 (1971).
14. Pauliny-Toth, I. I. K. et al. *Astr. J.* **86**, 371–385 (1981).
15. Schraml, J. et al. *Astrophys. J. Lett.* **251**, L57–L60 (1981).
16. Spencer, J. H., Johnston, K. J., Pauliny-Toth, I. I. K. & Witzel, A. *Astrophys. J. Lett.* **251**, L61–L64 (1981).
17. Unwin, S. C. *IAU Symp. No. 97*, 357–358 (1982).
18. Baath, L. B. et al. *Astrophys. J. Lett.* **243**, L123–L126 (1981).
19. Wilkinson, P. N., Readhead, A. C. S., Anderson, B. & Purcell, G. H. *Astrophys. J.* **232**, 365–381 (1979).
20. Browne, I. W. A. et al. *Mon. Not. R. astr. Soc.* **198**, 673–688 (1982).
21. Mutel, R. L., Allen, H. D. & Phillips, R. B. *Nature* **294**, 236–238 (1981).
22. Phillips, R. B. & Mutel, R. L. *Astrophys. J. Lett.* **257**, L19–L21 (1982).
23. Joshi, M. N. & Gopal-Krishna, *Mon. Not. R. astr. Soc.* **178**, 717–734 (1977).

24. Brown, R. L. et al. *Astrophys. Lett.* **21**, 105–110 (1981).
25. Fanti, R., Ficarra, A., Mantovani, F., Padrielli, L. & Weiler, K. *Astr. Astrophys. Suppl.* **36**, 359–369 (1979).
26. Fisher, J. R. & Erickson, W. C. *Astrophys. J.* **242**, 884–893 (1980).
27. Pearson, T. J., Readhead, A. C. S. & Wilkinson, P. N. *Astrophys. J.* **236**, 714–723 (1980).
28. Cotton, W. D., Geldzahler, B. J. & Shapiro, I. I. *IAU Symp. No. 97*, 301–303 (1982).
29. Perley, R. A. *Optical Jets in Galaxies* 77–81 (ESO/ESA Workshop 1981).
30. Marscher, A. P. & Scott, J. S. *Publ. astr. Soc. Pacif.* **92**, 127–133 (1980).
31. Perley, R. A., Fomalont, E. B. & Johnston, K. J. *Astr. J.* **85**, 649–658 (1980).
32. Bridle, A. H. *IAU Symp. No. 97*, 121–128 (1982).
33. Miley, G. K. A. *Rev. Astr. Astrophys.* **18**, 165–218 (1980).
34. Linfield, R. *Astrophys. J.* **254**, 465–471 (1982).
35. Scheuer, P. A. G. & Readhead, A. C. S. *Nature* **277**, 182–185 (1979).
36. van Groningen, E., Miley, G. K. & Norman, C. A. *Astr. Astrophys.* **90**, L7–L9 (1980).
37. Macklin, J. T. *Mon. Not. R. astr. Soc.* **196**, 967–986 (1981).
38. Hooley, A., Longair, M. S. & Riley, J. M. *Mon. Not. R. astr. Soc.* **182**, 127–145 (1978).
39. Jägers, W. J., van Brugel, W. J. M., Miley, G. K., Schilizzi, R. T. & Conway, R. G. *Astr. Astrophys.* **105**, 278–283 (1982).
40. Moore, P. K., Browne, I. W. A., Daintree, E. J., Noble, R. G. & Walsh, D. *Mon. Not. R. astr. Soc.* **197**, 325–338 (1981).
41. Sandage, A. & Tammann, G. A. *Astrophys. J.* **256**, 339–345 (1982).
42. Simard-Normandin, M., Kronberg, P. P. & Button, S. *Astrophys. J. Suppl.* **45**, 97–111 (1981).
43. Davis, R. J., Stannard, D. & Conway, R. G. *Mon. Not. R. astr. Soc.* **185**, 435–440 (1978).
44. Baath, L. B. et al. *Astr. Astrophys.* **86**, 364–372 (1980).
45. Readhead, A. C. S., Cohen, M. H., Pearson, T. J. & Wilkinson, P. N. *Nature* **276**, 768–771 (1978).

Primary structure of α -subunit precursor of *Torpedo californica* acetylcholine receptor deduced from cDNA sequence

Masaharu Noda*, Hideo Takahashi*, Tsutomu Tanabe*, Mitsuyoshi Toyosato*, Yasuji Furutani*, Tadaaki Hirose†, Michiko Asai†, Seiichi Inayama†, Takashi Miyata† & Shosaku Numa*

* Department of Medical Chemistry, Kyoto University Faculty of Medicine, Kyoto 606, Japan

† Pharmaceutical Institute, Keio University School of Medicine, Tokyo 160, Japan

‡ Department of Biology, Kyushu University Faculty of Science, Fukuoka 812, Japan

DNA sequences complementary to the Torpedo californica electroplax mRNA coding for the α -subunit precursor of the acetylcholine receptor were cloned. The nucleotide sequence of the cloned cDNA indicates that the precursor consists of 461 amino acids including a prepeptide of 24 amino acids. Possible sites for acetylcholine binding and antigenic determinants on the α -subunit molecule are discussed.

THE nicotinic acetylcholine receptor that mediates synaptic transmission at the vertebrate neuromuscular junction is the best characterized neurotransmitter receptor from both physiological and biochemical viewpoints (reviewed in refs 1, 2). The acetylcholine receptor from the electroplax of the electric ray *Torpedo californica* is composed of five subunits present in a molar stoichiometry of $\alpha_2\beta\gamma\delta$ (refs 3–5); the apparent molecular weights of the α -, β -, γ - and δ -subunits are 40,000, 50,000, 60,000 and 65,000, respectively^{6–10}. The pentameric protein complex is responsible for the binding of agonists,

cholinergic antagonists, α -toxins, histrionicotoxin and local anaesthetics as well as for cation translocation (reviewed in refs 1, 2). The α -subunit contains part or all of the acetylcholine binding site^{6,11,12}. A prerequisite for the understanding of the molecular basis for these various functions is to elucidate the primary structures of the constituent polypeptides of the receptor. Recently, the amino-terminal 54–56 amino acids of all four polypeptides derived from *T. californica* have been sequenced by Raftery *et al.*⁵, who have found considerable amino acid homology between the subunits, suggesting a common genetic

origin during the evolution of the acetylcholine receptor.

We have now applied recombinant DNA techniques to study the primary structure of the α -subunit of the *T. californica* acetylcholine receptor. DNA sequences complementary to the mRNA coding for this polypeptide have been cloned, and nucleotide sequence analysis of the cloned cDNA has revealed the whole amino acid sequence of the α -subunit precursor. Some structural features of the α -subunit molecule are discussed with respect to acetylcholine binding and immunogenicity.

Isolation of cDNA clones and strategy of DNA sequencing

The approach used to clone cDNA for the α -subunit precursor of the acetylcholine receptor was essentially the same as those described for cloning cDNA for preproenkephalin A (ref. 13) and preproenkephalin B (ref. 14). A library of cDNA clones derived from poly(A)-containing RNA from an electric organ of *T. californica* was constructed using the plasmid DNA vector

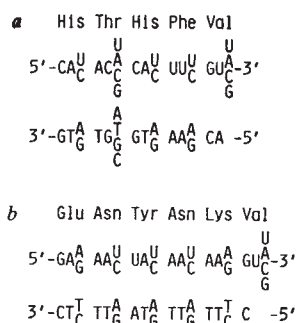


Fig. 1 Synthetic oligodeoxyribonucleotides used for probing cloned cDNA for the α -subunit precursor of the acetylcholine receptor. All possible coding sequences and the corresponding tetradecamers or hexadecamers used are given for the pentapeptide composed of residues 25–29 (a) and the hexapeptide composed of residues 13–18 (b); for the numbering of amino acid residues, see Fig. 3. Probe a was synthesized as two pools that contained A/T or G/C at the position corresponding to the third nucleotide of the Thr codon. Probe b was synthesized as two pools that contained A or G at the position corresponding to the third nucleotide of the left Asn codon (residue 14). For screening cDNA clones, an equimolar mixture of two pools each was used.

elaborated by Okayama and Berg¹⁵. The cDNA library was screened by hybridization with a mixture of 32 oligodeoxyribonucleotides (Fig. 1a) synthesized by the triester method¹⁶. These tetradecamers represented all possible cDNA sequences predicted from the pentapeptide sequence His-Thr-His-Phe-Val (residues 25–29) contained in the amino-terminal region of the α -subunit⁵ (excluding the third nucleotide residue of the Val codon).

Fifty-seven hybridization-positive clones were isolated from $\sim 2 \times 10^5$ transformants and were re-screened by hybridization with a second mixture of 32 oligodeoxyribonucleotides (Fig. 1b) synthesized similarly¹⁶. These hexadecamers corresponded to all possible cDNA sequences predicted from the hexapeptide sequence Glu-Asn-Tyr-Asn-Lys-Val (residues 13–18) present also in the amino-terminal region of the α -subunit⁵ (excluding the second and third nucleotide residues of the Val codon). Thus, 20 clones that hybridized with both probes were isolated. Restriction mapping of these clones with the endonucleases *Pst*I, *Pvu*II, *Rsa*I and *Hin*fI showed the presence of common restriction sites in all clones, suggesting that they represented a single mRNA species. Two of these clones, pACR α 30 and pACR α 44 (carrying the apparently largest cDNA insert), were subjected to nucleotide sequence analysis by the procedure of Maxam and Gilbert¹⁷ according to the strategy indicated in Fig. 2 (see Fig. 2 legend for experimental details of cDNA cloning).

Nucleotide sequence of mRNA and characteristics of the protein sequence deduced

The primary structure of the *T. californica* mRNA encoding the α -subunit precursor of the acetylcholine receptor (Fig. 3) was deduced from the 2,045-nucleotide sequence (excluding the poly(dA)·poly(dT) tract and poly(dG)·poly(dC) tail) determined for the cDNA inserts of clones pACR α 30 and pACR α 44. The nucleotide sequence of the whole protein-coding region was determined with both clones. No nucleotide difference was observed between the sequences determined for the two clones, apparently because the template RNA used for *in vitro* cDNA synthesis was derived from a single electric organ.

The sequence of nucleotide residues 1–162 corresponds precisely to the sequence determined for the amino-terminal 54 amino acids of the acetylcholine receptor α -subunit⁵. The amino acid sequence of the cryptic portion of the α -subunit was

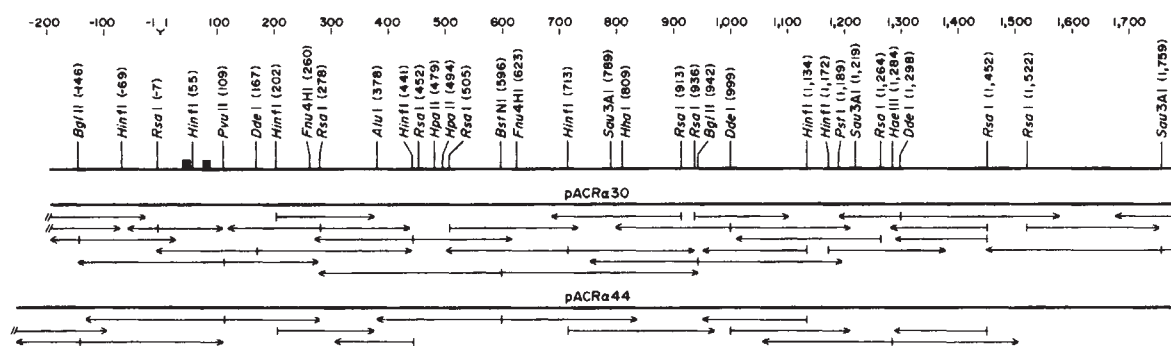
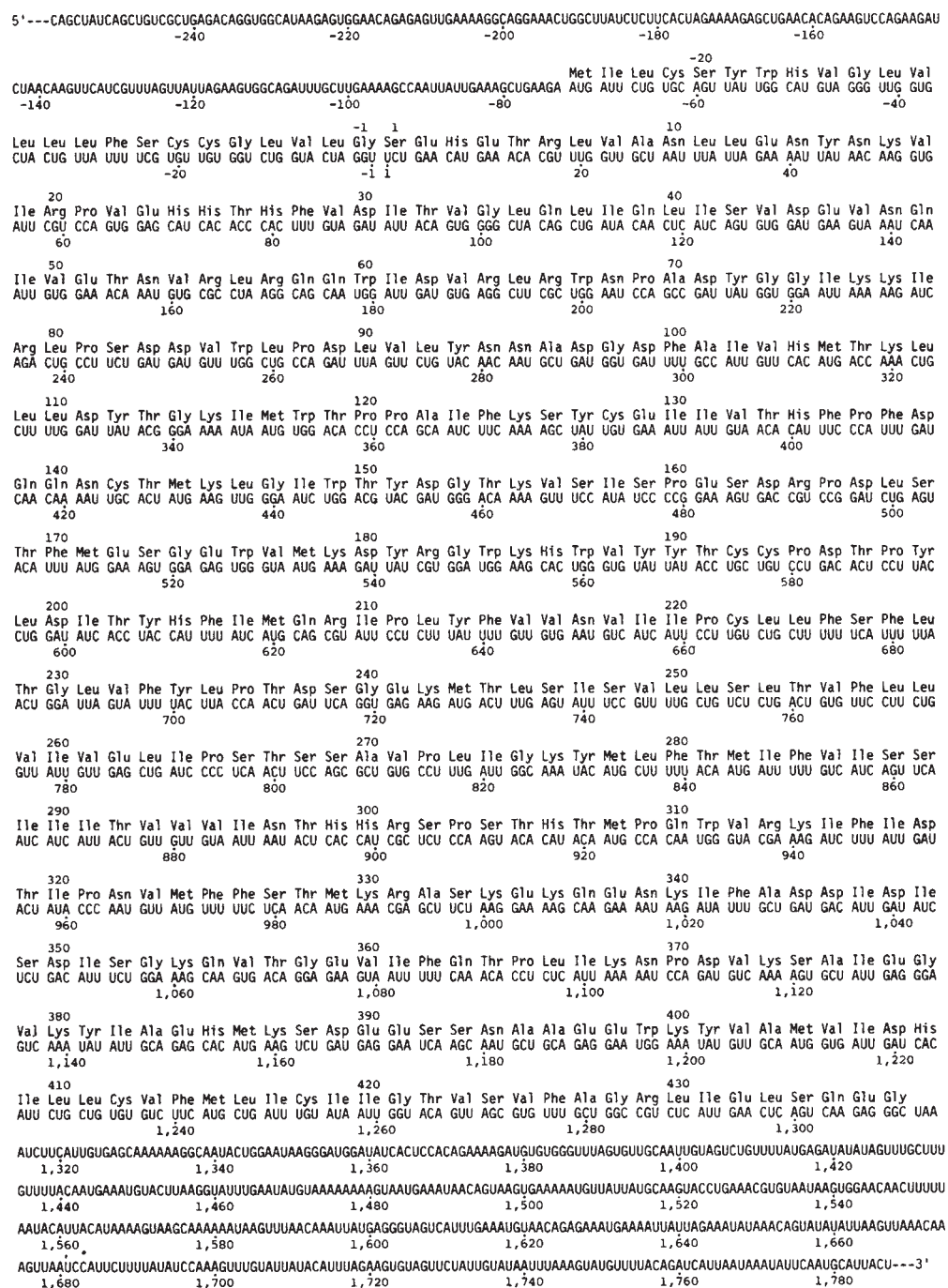


Fig. 2 Strategy for sequencing the cDNA inserts in clones pACR α 30 and pACR α 44. For the isolation of the clones, a cDNA library was constructed by the method of Okayama and Berg¹⁵, using 2.3 μ g of poly(A) RNA from an electric organ of *T. californica* (Pacific Bio-Marine Laboratories) and 1.4 μ g of the vector-primer DNA; avian myeloblastosis virus reverse transcriptase was provided by Dr J.W. Beard. Poly(A) RNA was isolated by subjecting the total RNA extracted³⁷ to oligo(dT)-cellulose chromatography³⁸. *Escherichia coli* χ 1776 (ref. 39) or HB101 (ref. 40) was used for transformation⁴¹, and ampicillin-resistant transformants were screened⁴² by hybridization at 40 °C with the oligodeoxyribonucleotide probe a (Fig. 1a) labelled with ³²P at the 5' end. The hybridization-positive clones were re-screened with the oligodeoxyribonucleotide probe b (Fig. 1b) as mentioned above. The restriction map displays only relevant restriction endonuclease sites, which are identified by numbers indicating the 5'-terminal nucleotide generated by cleavage (for the nucleotide numbers, see Fig. 3). The poly(dA)·poly(dT) tract and poly(dG)·poly(dC) tail are not included in the restriction map. The sequences corresponding to the hybridization probes used are indicated by closed boxes. DNA sequencing was carried out by the procedure of Maxam and Gilbert¹⁷. The direction and extent of sequence determinations are shown by horizontal arrows under each clone used; the sites of 5'-end labelling are indicated by short vertical lines at the end of arrows. The slash marks at the end of arrows mean that the site of 5'-end labelling was located on the vector DNA^{15,43} as follows (the numbers in parentheses indicate the distance in nucleotides between the site of labelling and the *Pst*I site (on the 5' side) or *Pvu*II site (on the 3' side) flanking the cDNA insert): *Bst*NI (left upper mark) (45), *Rsa*I (left lower mark) (22) and *Pvu*II (right mark) (0) for pACR α 30; *Bst*NI (45) for pACR α 44. Further details of the sequencing procedures have been described previously⁴⁴.

Fig. 3 Primary structure of the *T. californica* mRNA coding for the α -subunit precursor of the acetylcholine receptor. The nucleotide sequence of the mRNA was deduced from those of the cDNA inserts in clones pACR α 30 and pACR α 44. Nucleotide residues are numbered in the 5' to 3' direction, beginning with the first residue of the codon specifying the amino-terminal serine of the mature α -subunit, and the nucleotides on the 5' side of residue 1 are indicated by negative numbers. The 5'-terminal sequence presented does not extend to the 5' end of the mRNA (see text), whereas the 3'-terminal sequence shown is followed by a poly(A) tract of ~70 nucleotides connected with the vector DNA sequence^{15,43,45}, thus representing the complete sequence of this region. The predicted amino acid sequence is displayed above the nucleotide sequence, and amino acid residues are numbered beginning with the amino-terminal serine of the mature α -subunit.



prepeptide of the α -subunit precursor. Thus, it seems likely that this hydrophobic sequence is somehow involved in the translocation of the acetylcholine receptor subunit into the cytoplasmic membrane.

The amino acid composition of the *T. californica* acetylcholine receptor α -subunit predicted from the cDNA sequence agrees fairly well with that determined^{4,10}. However, the predicted molecular weight of the α -subunit (50,116) is considerably larger than the values reported, which range between 39,000 and 48,000 depending on the electrophoresis system and the method of evaluation⁶⁻¹⁰. It is generally recognized that molecular weights estimated by SDS-polyacrylamide gel electrophoresis may be quite inaccurate (for example, see ref. 2). This may hold especially for polypeptides derived from membrane proteins that have higher capacities to bind the detergent due to their larger contents of hydrophobic amino acids²¹⁻²⁴. On the other hand, the possibility that a carboxy-terminal peptide is eliminated by post-translational cleavage cannot be excluded either. Note in this context that the paired basic residues Lys-Arg and Arg-Lys, which generally represent the sites of proteolytic processing^{13,20}, are present at positions 330-331 and 313-314, respectively. Computer analysis of the nucleotide and the amino acid sequence of the α -subunit precursor has not detected such internal homology units as observed for the immunoglobulin polypeptides²⁵.

It has been reported that each subunit of the *T. californica* acetylcholine receptor contains carbohydrate^{4,7,10,11}. The only possible site of *N*-glycosidic linkage in the α -subunit molecule is the asparagine residue in the sequence Asn-Cys-Thr at positions 141-143 (ref. 26). It is possible that *O*-glycosidic carbohydrate units are also present in the α -subunit, because the acetylcholine receptor is known to contain *O*-substituted serine and threonine residues which are probably glycosylated¹⁰.

The α -subunit of the *T. californica* acetylcholine receptor contains seven cysteine residues (excluding the three cysteine residues in the prepeptide). This is consistent with the presence of free sulphhydryls in the receptor molecule^{27,28}. There is evidence that at least two of the cysteine residues in the α -subunit form a disulphide bridge near the acetylcholine binding site^{2,29}. This disulphide can be reduced and alkylated by affinity reagents. Such affinity-labelling studies suggest that, in the acetylcholine binding region, the distance between the negative subsite binding the quaternary ammonium group and one of the sulphhydryl groups formed by reduction is about 1 nm (refs



Fig. 4 Autoradiogram of blot hybridization analysis of *T. californica* electroplax RNA with a cDNA probe. Poly(A) RNA (10 μ g) was denatured with 1 M glyoxal and 50% dimethyl sulphoxide⁴⁶, electrophoresed on a 1.5% agarose gel and transferred to diazobenzoyloxymethyl paper⁴⁷. The hybridization probe used was the 1,088-base pair *Bgl*II fragment (residues -146 to 942) derived from clone pACR α 30; the probe was labelled by nick-translation⁴⁸ with [α -³²P]dCTP. The size markers used were *E. coli* rRNA⁴⁶.

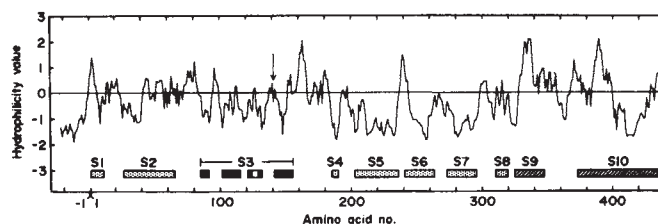


Fig. 5 Predicted secondary structures and hydrophilicity profile of the α -subunit precursor of the acetylcholine receptor. The positions of the predicted structures of α -helix and/or β -sheet (S1-S10) are indicated as follows: shaded boxes, probably α -helix; closed boxes, probably β -sheet; stippled boxes, α -helix or β -sheet; the β -sheet structure corresponding to the open portion in the third box of S3 may partially break. Secondary structure predictions for the prepeptide region are not included. The averaged hydrophilicity value of a hexapeptide composed of amino acid residues $i-2$ to $i+3$ has been plotted against i , where i represents amino acid number (see Fig. 3). The arrow indicates the position of the asparagine residue (residue 141) as a possible site of glycosylation. For the procedures of computer analysis, see refs 33 and 34.

2, 29). In the native unreduced state, a second subsite that is involved in the binding of acetylcholine through hydrophobic interaction or hydrogen bonding is postulated^{2,29}. Because three of the seven cysteine residues present in the α -subunit molecule, that is, residues 222, 412 and 418, are located in predicted secondary structures of strong hydrophobicity (see Fig. 5), it seems improbable that they appear on the outer surface of the membrane or receptor molecule for ready access to affinity reagents. At least two of the remaining four cysteine residues (residues 128, 142, 192 and 193) may take part in the disulphide bridge present in close proximity to the acetylcholine binding site. (The formation of a disulphide bond between the adjacent cysteine residues 192 and 193 is unlikely³⁰.) Because the asparagine residue at position 141 is most probably glycosylated (see above), the region surrounding it would reside on the extracellular surface of the receptor molecule. Moreover, this region encompasses a predicted β -turn structure lying between predicted β -sheet structures which may be tightly packed (see Fig. 5), so that its conformation would be more or less stabilized.

The nucleotide binding site of dehydrogenases occurs at an end of a pleated sheet structure on the outer surface of the molecules³¹. A similar conformation is observed in the hyper-variable regions of the immunoglobulin molecules³². Thus, it is tempting to assume that the acetylcholine binding site is located in the region including the cysteine residues at positions 128 and 142. Measurements on Corey-Pauling-Koltun space-filling models are consistent with the hypothesis that the aspartic acid residue at position 138 (or the glutamic acid residue at position 129) as the negative subsite, together with, for example, the histidine residue at position 134 or a hydrophobic residue in its vicinity as the second subsite, may be involved in the binding of acetylcholine. An alternative possibility is that the aspartic acid residue at position 195 or 200 near the cysteine residues at positions 192 and 193 may act as the negative subsite; the region including these residues also contains a predicted β -turn structure surrounded by predicted secondary structures (see Fig. 5).

The amino acid sequence of the acetylcholine receptor α -subunit precursor was analysed by computer for secondary structure and local hydrophilicity according to the procedures of Chou and Fasman³³ and of Hopp and Woods³⁴, respectively. Figure 5 shows the hydrophilicity profile along the polypeptide chain as well as the predicted structures of α -helix and/or β -sheet which are schematically represented by boxes (S1-S10). The segments that can form secondary structures extend over a wide range of the α -subunit molecule. The structure S10, most likely to be an α -helix, is particularly long (64 amino acid residues) and contains a large hydrophobic region in the middle, probably representing a transmembrane segment. The

structures S5, S6 and S7 also exhibit strong hydrophobicity, presumably being buried in the membrane or being involved in hydrophobic inter-subunit interaction. The structure S3 is composed of a set of four β -sheets which may be closely packed in an antiparallel orientation. This region encompasses a putative acetylcholine binding site as discussed above. The pleated sheet structure would facilitate the generation of a specific conformation for binding the neurotransmitter.

Hopp and Woods³⁴ have reported that the point of highest local hydrophilicity is invariably located in or immediately adjacent to a protein antigenic determinant. This allows us to predict possible candidates for antigenic determinants on the α -subunit molecule, that is, the peptide segments composed of residues 161–166, 330–340 and 387–392. It has been observed that some antigenic determinants are not correlated with hydrophilicity, although there seems to be a correlation of many antigenic determinants with local upspikes of the hydrophilicity profile³⁴. Thus, for example, the residues 1–6 and 237–242 may

also represent possible antigenic determinants. Knowledge of the antigenic determinants on the acetylcholine receptor and its subunits is important because antibodies directed to them are the primary pathological agents in myasthenia gravis and its animal model, experimental autoimmune myasthenia gravis (reviewed in ref. 35). It has been reported that most of the antibodies produced by human patients as well as by experimental animals are directed to the 'main immunogenic region' which is located on the extracellular surface of the α -subunit and is distinct from the acetylcholine binding site³⁶. The possible acetylcholine binding site and antigenic determinants we have proposed are consistent with this finding.

We thank Dr Paul Berg and Dr Hiroto Okayama for providing us with their high-efficiency cloning system, and Dr Tadashi Tanabe for helpful discussions. This investigation was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases.

Received 20 July; accepted 3 September 1982.

- Heidmann, T. & Changeux, J. P. *A. Rev. Biochem.* **47**, 317–357 (1978).
- Karlin, A. in *Cell Surface Reviews* Vol. 6 (eds Cotman, C. W., Poste, G. & Nicolson, G. L.) 191–260 (North-Holland, Amsterdam, 1980).
- Reynolds, J. & Karlin, A. *Biochemistry* **17**, 2035–2038 (1978).
- Lindstrom, J., Merlie, J. & Yogeewaran, G. *Biochemistry* **18**, 4465–4470 (1979).
- Rafferty, M. A., Hunkapiller, M. W., Strader, C. D. & Hood, L. E. *Science* **208**, 1454–1456 (1980).
- Weill, C. L., McNamee, M. G. & Karlin, A. *Biochem. biophys. Res. Commun.* **61**, 997–1003 (1974).
- Rafferty, M. A., Vandlen, R. L., Reed, K. L. & Lee, T. *Cold Spring Harb. Symp. quant. Biol.* **40**, 193–202 (1975).
- Hucho, F., Bandini, G. & Suárez-Isla, B. A. *Eur. J. Biochem.* **83**, 335–340 (1978).
- Froehner, S. C. & Rafto, S. *Biochemistry* **18**, 301–307 (1979).
- Vandlen, R. L., Wu, W. C. S., Eisenach, J. C. & Rafferty, M. A. *Biochemistry* **18**, 1845–1854 (1979).
- Karlin, A., Weill, C. L., McNamee, M. G. & Valderrama, R. *Cold Spring Harb. Symp. quant. Biol.* **40**, 203–213 (1975).
- Damle, V. & Karlin, A. *Biochemistry* **17**, 2039–2045 (1978).
- Noda, M. *et al. Nature* **295**, 202–206 (1982).
- Kakidani, H. *et al. Nature* **298**, 245–249 (1982).
- Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161–170 (1982).
- Ito, H., Ike, Y., Ikuta, S. & Itakura, K. *Nucleic Acids Res.* **10**, 1755–1769 (1982).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
- Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
- Blöbel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 852–862 (1975).
- Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J. & Tager, H. S. *Ann. N.Y. Acad. Sci.* **343**, 1–16 (1980).
- Grefrath, S. P. & Reynolds, J. A. *Proc. natn. Acad. Sci. U.S.A.* **71**, 3913–3916 (1974).
- Robinson, N. C. & Tanford, C. *Biochemistry* **14**, 369–378 (1975).
- Frank, R. N. & Rodbard, D. *Archs Biochem. Biophys.* **171**, 1–13 (1975).
- Miyake, J., Ochiai-Yanagi, S., Kasumi, T. & Takagi, T. *J. Biochem.* **83**, 1679–1686 (1978).
- Hill, R. L., Delaney, R., Fellows, R. E. Jr & Lebovitz, H. E. *Proc. natn. Acad. Sci. U.S.A.* **56**, 1762–1769 (1966).
- Marshall, R. D. *Biochem. Soc. Symp.* **40**, 17–26 (1974).
- Eldefrawi, M. E., Eldefrawi, A. T. & Wilson, D. B. *Biochemistry* **14**, 4304–4310 (1975).
- Chang, H. W. & Bock, E. *Biochemistry* **16**, 4513–4520 (1977).
- Karlin, A. *J. gen. Physiol.* **54**, 245s–264s (1969).
- Dayhoff, M. D. in *Atlas of Protein Sequence and Structure* Vol. 5, 77–268 (National Biomedical Research Foundation, Washington DC, 1976).
- Liljas, A. & Rossmann, M. G. *A. Rev. Biochem.* **43**, 475–507 (1974).
- Poljak, R. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **70**, 3305–3310 (1973).
- Chou, P. Y. & Fasman, G. D. *A. Rev. Biochem.* **47**, 251–276 (1978).
- Hopp, T. P. & Woods, K. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 3824–3828 (1981).
- Lindstrom, J. & Dau, P. A. *Rev. Pharmac. Tox.* **20**, 337–362 (1980).
- Tzartos, S. J., Seybold, M. & Lindstrom, J. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 188–192 (1982).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
- Curtiss, R. III *et al. in Recombinant Molecules, Impact on Science and Society* (eds Beers, R. F. & Bassett, E. G.) 45–56 (Raven, New York, 1977).
- Boyer, H. W. & Roulland-Dussoix, D. *J. molec. Biol.* **41**, 459–472 (1969).
- Wahl, G. M., Padgett, R. A. & Stark, G. R. *J. biol. Chem.* **254**, 8679–8689 (1979).
- Hanahan, D. & Meselson, M. *Gene* **10**, 63–67 (1980).
- Reddy, V. B. *et al. Science* **200**, 494–502 (1978).
- Nakanishi, S. *et al. Nature* **278**, 423–427 (1979).
- Sutcliffe, J. G. *Nucleic Acids Res.* **5**, 2721–2728 (1978).
- McMaster, G. K. & Carmichael, G. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835–4838 (1977).
- Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350–5354 (1977).
- Weinstock, R., Sweet, R., Weiss, M., Cedar, H. & Axel, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1299–1303 (1978).

Human metallothionein genes—primary structure of the metallothionein-II gene and a related processed gene

Michael Karin* & Robert I. Richards†

* Metabolic Research Unit, Department of Medicine, Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA and Department of Microbiology, University of Southern California School of Medicine, 2025 Zonal Avenue, HMR-401, Los Angeles, California 90033, USA

† Center for Recombinant DNA Research and Department of Genetics, Research School of Biological Sciences, Australian National University, Canberra, ACT 2601, Australia

The complete nucleotide sequence of two of the human metallothionein gene family has been compared. One is a functional metallothionein-II gene, the other a pseudogene, lacking introns, terminating in a poly(A) tail and flanked by two direct repeats. In addition, we have detected a size polymorphism associated with the processed gene in the population examined, and we have observed a region of apparent secondary structure homology between a 5' flanking region of the functional metallothionein-II gene and that of a mouse metallothionein-I gene.

METALLOTHIONEINS (MTs) are a family of low-molecular weight heavy metal binding proteins, unique in their high cysteine content. MTs specifically bind heavy metals, such as zinc, cadmium, copper and mercury. MTs and related heavy metal binding proteins are ubiquitously distributed in the animal and plant kingdoms and are even found in prokaryotes¹. These proteins are inducible in experimental animals and cultured

cells both by exposure to heavy metal ions and glucocorticoid hormone². Their ubiquitous distribution and high inducibility make them an attractive model system to study regulation of gene expression.

Human MTs have been isolated from liver³ and cultured cells^{4,5} where they are present in two major electrophoretically distinguishable forms known as MT-I and MT-II³. The amino

acid sequence of both human MT-I and MT-II have been determined^{1,6}. Although MT-II seems to be a unique protein, MT-I is a mixture of at least three different proteins generating a heterogeneous amino acid sequence.

We have been studying the regulation of MT synthesis in HeLa cells and have shown that either heavy metals or glucocorticoids are primary inducers of MT mRNA⁷. As neither one of them affects the stability of the mRNA⁸, transcriptional control has been suggested. Recently, Palmiter and co-workers have directly demonstrated that either heavy metals or glucocorticoids increase the transcription rate of the mouse MT-I gene^{9,10}. To define the components involved in the molecular mechanisms responsible for the regulation of MT gene expression in man, we isolated and characterized DNA fragments containing the different genes.

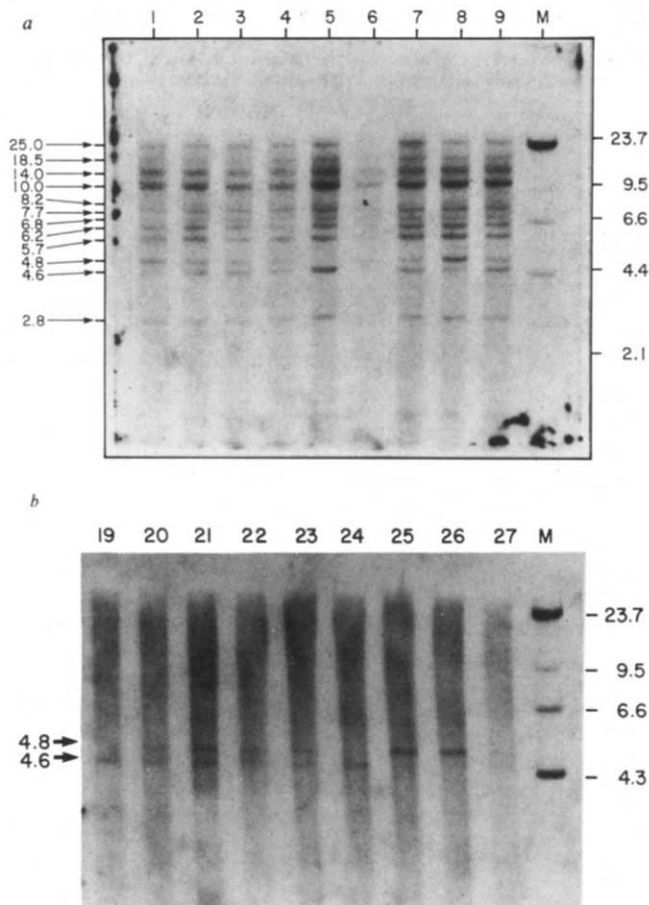


Fig. 1 Detection of MT genes in human genomic DNA. Human genomic DNA was prepared from white blood cells, digested with *EcoRI*, the fragments were separated by electrophoresis on 1% agarose gels and transferred to nitrocellulose filters as described by Bell *et al.*³³. *a*, The *BamHI*-*PvuII* coding region of phMT-II₃ (ref. 11) was labelled by nick-translation³⁴ using both [α -³²P]dATP and [α -³²P]dCTP. Hybridization was as described by Wahl *et al.*³⁵, with slight modifications. We have included an additional washing step, using 50% formamide, 5×SSC at 42°C for 60 min, before the final wash step, using 0.1×SSC at 42°C was performed (1×SSC=0.15 M NaCl, 0.015 M Na citrate). Individuals are indicated by number above the autoradiogram, the sizes of the hybridizing fragments are on the left and the sizes of the bacteriophage λ *HindIII* fragments, which served as molecular weight markers, are on the right. *b*, To demonstrate more clearly the size polymorphism associated with the MT-II_B pseudogene, an hybridization probe specific for that gene was used. The *EcoRI*-*BamHI* fragment containing the 5' portion of the MT-II_B gene (see Fig. 2) was labelled by nick-translation³⁴. Hybridization to the blots of *EcoRI*-digested human DNA was as described above. Final washing, using 0.1×SSC, was performed at 60°C.

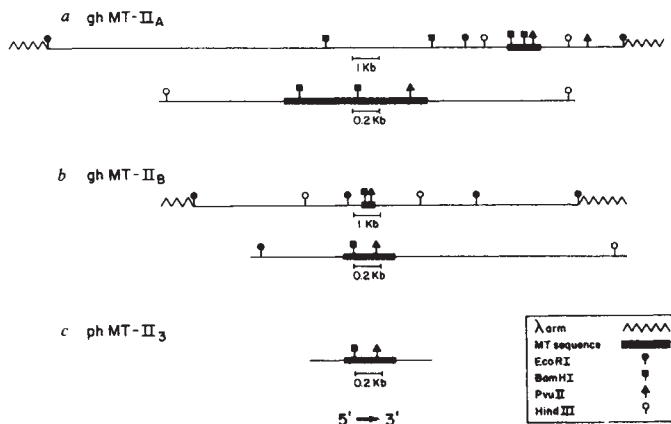


Fig. 2 Restriction maps of the two human MT-II genes. About 1.5×10^6 plaques (10 genome equivalents) of the human genomic library prepared by Lawn *et al.*¹² were screened by the method of Benton and Davis¹⁴ using the *BamHI*-*PvuII* coding region of phMT-II₃ as a hybridization probe. Of 200 positive plaques detected, 100 were subjected to three cycles of plaque purification which finally resulted in 53 positive clones. Phage DNA was prepared, digested with *EcoRI*, *BamHI* or *PvuII*, the fragments were separated on 1% agarose gels, transferred to nitrocellulose¹³ and the fragments containing MT structural sequences were identified by hybridization to the coding region probe. The genes coding for MT-II were finally identified using the 5' and 3' untranslated regions of phMT-II₃ as hybridization probes. Shown are restriction maps of the largest Charon 4A clones isolated and of subclones of the MT-II_A (*a*) and MT-II_B (*b*) genes. A restriction map of phMT-II₃ (*c*), the cDNA clone coding for MT-II mRNA¹¹, is also shown for comparison.

As described previously¹¹, cDNA clones for the different human MT genes were isolated. Using one of these clones, we screened a human genomic DNA library¹² and isolated phage λ clones containing most of the different MT genes. Here we describe the nucleotide sequence of a functional gene coding for human MT-II and a closely related processed pseudogene.

Metallothioneins are encoded by a large gene family

Previously we obtained a full-length cDNA clone of human MT-II mRNA, phMT-II₃ (ref. 11). The cDNA can be conveniently divided into: (1) 5' untranslated region, from the 5' end of the cDNA to the *BamHI* site; (2) coding region, from the *BamHI* site to the *PvuII* site; and (3), 3' untranslated region, from the *PvuII* site to the 3' end of the cDNA (see Fig. 2). To estimate the number and the sizes of the MT genes in the human genome we have used the *BamHI*-*PvuII* coding region as a probe, since it is the region most likely to be conserved among the different genes^{1,6}. Human genomic DNA was digested with *EcoRI*, fractionated by electrophoresis on agarose gels and transferred to nitrocellulose filters¹³. At least 11 or 12 different bands hybridizing to the MT-II coding region probe could be detected (Fig. 1). About 15 hybridizing bands can be detected when genomic DNA is digested with *HindIII* (data not shown), suggesting a multiplicity of genes bearing at least 80% homology to the coding region of MT-II. One of the restriction fragments exhibits polymorphism and appears as either a 4.8- or 4.6-kilobase (kb) *EcoRI* band. This fragment corresponds to the MT-II_B pseudogene. The significance of this result is discussed below.

Isolation of the human MT genes

To isolate all the human MT genes, we screened a human genomic library¹² using the *BamHI*-*PvuII* coding region probe. After screening 1.5×10^6 recombinant phage, 200 positive plaques were detected; 100 of the hybrids were subjected to three cycles of plaque purification¹⁴ which finally yielded 53

positive clones containing fragments hybridizing with the MT-II coding region.

DNA was prepared from each of the 53 positive clones and was analysed by *EcoRI* digestion and Southern blotting. At least 18 different *EcoRI* bands hybridizing to the MT-II coding region probe were detected among the different clones analysed. All but two of the *EcoRI* fragments detected on the genomic blots (Fig. 1) were represented among these bands. The two exceptions were the largest fragments migrating at 25 and 18.5 kb. While most of the λ clones contained a single positively hybridizing *EcoRI* fragment, some of them contained two such fragments. Those clones were found to contain several MT-I genes suggesting close linkage of this subfamily (data not shown). Taken together with the data shown in Fig. 1, the results of the cloning experiment suggest there are about 11 or 12 different human MT genes.

To identify the gene(s) coding for MT-II, the major MT to be expressed in HeLa cells and human liver⁵ *EcoRI*, *BamHI* and *PvuII* restriction enzyme digests of the 53 genomic clones were screened for hybridization with the untranslated region probes. Only three of the genomic clones contained fragments hybridizing to the 5' or 3' untranslated regions of MT-II cDNA. The 5' untranslated region probe hybridized to only two of the *EcoRI* fragments of genomic DNA, containing MT coding regions migrating at 5.8 and 4.8 or 4.6 kb (data not shown). While all the different genes share sequence homology at coding regions, they differ in their 5' and 3' untranslated regions.

Further restriction endonuclease mapping analysis revealed that two of the clones contained some identical DNA fragments and therefore overlapped. The gene contained in these two clones is referred to as ghMT-II_B, while the gene in the third unique clone is referred to as ghMT-II_A. Restriction maps are shown in Fig. 2. Interestingly, the distance between the *BamHI* and *PvuII* restriction sites in ghMT-II_B is approximately 170 base pairs (bp), the same as their relative location in the mRNA sequence¹¹. Since this fragment comprises most of the coding region, which in the mouse contains two intervening sequences¹⁵, it was apparent that no intervening sequences were expected in this gene. The distance between these two sites in ghMT-II_A mapped at 670 bp, suggesting the presence of intervening sequences.

Structure of the MT-II_A gene

We have determined the complete nucleotide sequences of the two genes. ghMT-II_A is the authentic gene coding for the human MT-II protein (Fig. 3). It spans about 900 bp and contains two introns. Both introns obey the GT-AG splicing rule¹⁶ and are located at exactly the same positions within the coding region as the introns in the mouse MT-I gene¹⁵.

The start of transcription of the ghMT-II_A gene was determined by a modification of the S₁ nuclease mapping technique¹⁷. (The experimental protocol is described in Fig. 4 legend.) Three nuclease-resistant bands can be detected corresponding to fragments starting at the *BamHI* site at position +75 and ending at the T at -3, A at +1 and A at +3 (Fig. 4a). (Similar patterns were obtained when S₁ nuclease was used, instead of mung bean nuclease.) The middle start site was determined as position +1, since it seems to be used more frequently.

The sequence of the mouse MT-I gene at the 5' end of the mRNA, GTCACCACGACT, is almost identical to the corresponding region in the human gene, except for two changes: T to C at position -2 and A to C at position 7. In view of this homology we suggest that the start of transcription of the mouse MT-I gene is also at the first A, and not at the G residue at -3, as determined by Glanville *et al.*¹⁵. This discrepancy could be due to strong steric hindrance by the mouse MT-I cap structure which prevents complete digestion by S₁ nuclease. This possibility has actually been raised by Glanville *et al.*¹⁵.

The expression of the mouse MT-I gene is regulated by heavy metals after introduction into human cells¹⁸ and the human

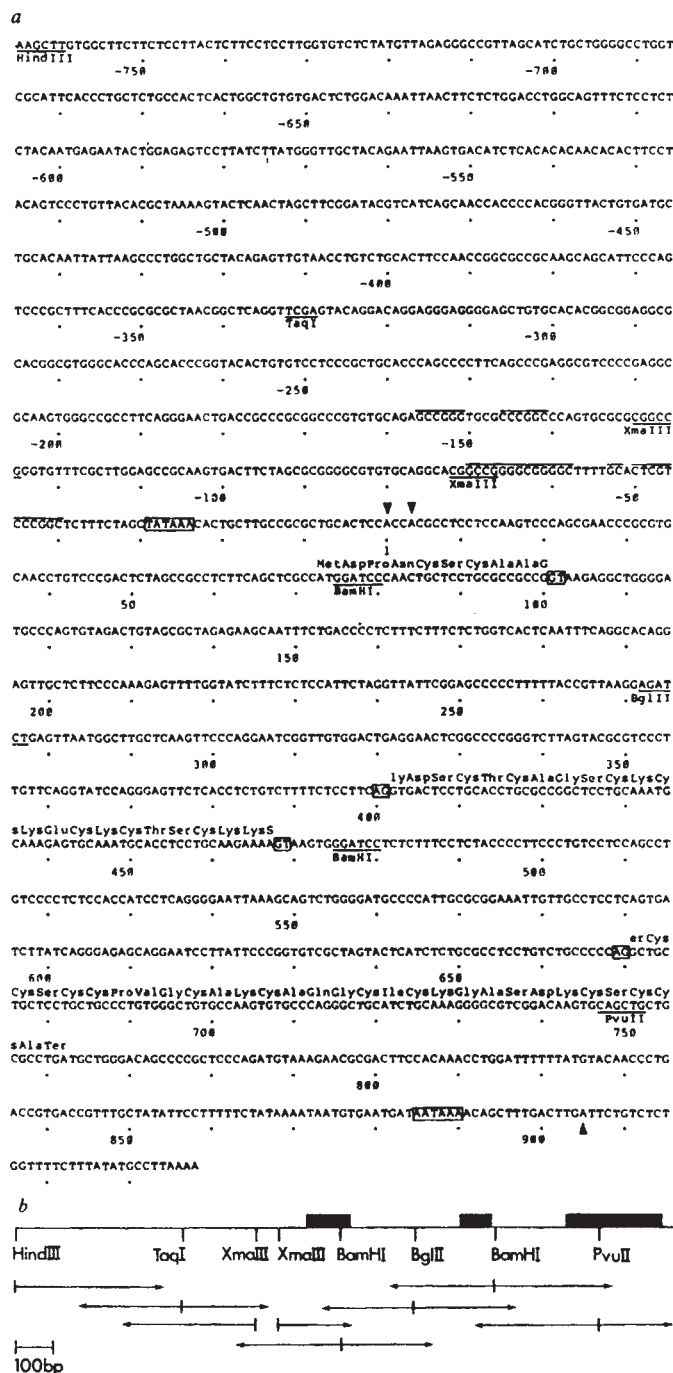


Fig. 3 Primary structure of the human MT-II_A gene. *a*, Nucleotide sequence of the MT-II_A gene. The major cap site (see Fig. 4) is numbered +1. Consensus sequences involved in transcription initiation (-30 to -24) and polyadenylation (886-891) are boxed. Introns run from nucleotides 100-401 and 469-671. The dyad symmetries in the 5' flanking region are underlined; ▼, ▲ indicate transcription initiation and polyadenylation sites, respectively. *b*, Sequencing strategy for the MT-II_A gene. Sequence analysis was performed on DNA isolated from a subclone in plasmid pBR322 (p84H) of the 3.2-kb *HindIII* restriction endonuclease fragment of ghMT-II_A (Fig. 2), by the chemical³⁶ and chain termination³⁷ methods. For the chemical analysis, restriction endonuclease-generated DNA fragments were end-labelled by end-filling³⁸ with [α -³²P]dNTPs (2,000-3,000 Ci mmol⁻¹; Amersham) and AMV reverse transcriptase (W. J. Beard, Life Sciences, Inc.). Modification and cleavage reactions were as described in ref. 36. Chain termination analysis was performed on restriction fragments cloned into bacteriophage M13 (strains mp8 and mp9)³⁹. Sequencing primer (5'-GTTTTCCTCAGTCACGAC-3') was given by Genentech, Inc. ■, Exons of the MT-II_A gene. Arrows indicate the direction and extent of sequencing (each sequence was determined at least twice).

MT-II_A gene is regulated in mouse cells (M.K., unpublished results). Based on the assumption that each of those genes is recognized by the regulatory factors from the other species, it was surprising not to find any extensive required homology in the immediate 5' flanking region of the two genes. Palmiter and co-workers have pointed out two regions of dyad symmetry upstream from the TATA box in the mouse MT-I gene¹⁴ and

suggested a possible regulatory role for these sequences. The regulatory role of the larger dyad symmetry (centred at position -55 on the MT-I gene) was supported by analysis of deletion mutants at the 5' flanking region of the mouse MT-I gene¹⁹. Deletions extending into this region or completely deleting it, abolish metal inducibility of the mouse gene. Close examination of the human MT-II_A gene at this region reveals similar dyad symmetry regions at almost identical positions to the mouse MT-I gene (Fig. 5).

Although of different composition, such a conserved structure may be of significance in the regulation of MT gene expression. We speculate that regulatory factors, most likely apothionein itself, recognize these secondary structures (possibly forming hairpin loops *in vivo*, as suggested in refs 20 and 21) and not a particular nucleotide sequence *per se*. We are presently constructing mutations at these sites to test this hypothesis.

Our previously obtained MT-II cDNA clone, phMT-II₃, did not end with a poly(A) tract¹¹. We have mapped the polyadenylation site directly by the nuclease protection technique¹⁷. Two major nuclease-resistant fragments could be detected on acrylamide-urea gels. Their sizes are 142 ± 2 and 153 ± 2 nucleotides (Fig. 4b). This would place the polyadenylation sites to the A residues at positions 893 and 905 (Fig. 3). It is possible that the observed band at position 893 is an artefact of the nuclease digestion method due to local melting of the hybrid at that region which is A+T-rich (a stretch of 11 A and T residues). The second termination site at 905 is located 15 nucleotides away from the end of the consensus sequence AATAAA, originally suggested by Proudfoot and Brownlee²², to serve as a polyadenylation signal.

Structure of the human MT-II_B gene

The sequence of the MT-II_B gene is shown in Fig. 6, in comparison with the structural sequence of the MT-II_A gene (without the introns). MT-II_B is highly homologous (95%) to an exact cDNA copy of MT-II mRNA. The homology to the MT-II_A gene starts exactly at the major cap site of the later gene and ends at its 3' end. At this point the MT-II_B pseudogene contains a stretch of 15 A residues, a landmark of RNA type processing. The MT-II_B gene is most likely derived from MT-II_A or its ancestor. The two genes share extensive homology in their 5' and 3' untranslated regions. As discussed earlier, the homology between the different MT genes is restricted to the coding region. (We have not detected any other MT-II-like genes among our genomic clones.)

The MT-II_B gene has accumulated several mutations, relative to the MT-II_A gene. Surprisingly, most of the mutations within its coding region are silent changes in the third codon position, only one causing an amino acid change of Gly to Ser. Since the reading frame is maintained and there are no premature termination codons, this gene does probably code for a functional protein although it is not expressed.

The MT-II_B pseudogene is flanked by two long, slightly imperfect, direct repeats which might have been generated by insertion site duplication²³. The 5' and 3' direct repeats are 27 and 25 nucleotides long, respectively, of which 21 nucleotides are identical. The 5' repeat does not flank the gene immediately as expected, but there is actually a trinucleotide spacer between

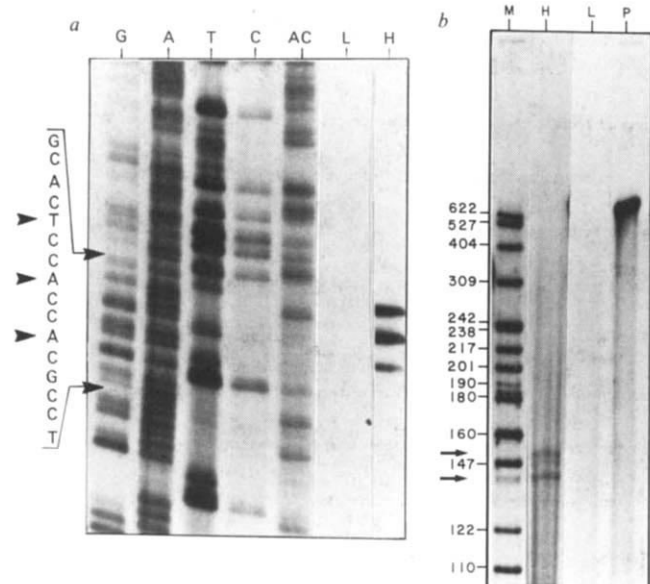


Fig. 4 Identification of the MT-II mRNA transcription start and termination sites. *a*, Plasmid ph84H, a *Hind*III subclone of the MT-II_A gene, was digested with *Bam*HI, 5' end-labelled using T₄ polynucleotide kinase (PL Biochemicals) and [γ -³²P]ATP and subjected to secondary digest with *Hind*III. The labelled *Hind*III-*Bam*HI fragment containing the 5' region of the MT-II_A gene was isolated by electrophoresis on 1% agarose gel and electroelution. Portions of the probe (50,000 c.p.m.) were co-precipitated with 100 μ g of total RNA from cadmium-induced HeLa (H) of LTK (L) cells, the precipitate was dissolved in 30 μ l of 80% formamide, 0.4 M NaCl, 40 mM PIPES pH 6.5, 1 mM EDTA, and overlaid with paraffin oil. The nucleic acids were denatured for 5 min at 85 °C and allowed to hybridize overnight at 54 °C. The sample was then diluted with 370 μ l of 0.25 M NaCl, 30 mM NaOAc pH 4.6, 1 mM ZnCl₂, 5% glycerol and incubated with 450 units of mung bean nuclease (PL Biochemicals) at 37 °C for 90 min. The digestion was terminated by phenol-CHCl₃ extraction and the nucleic acids were isolated by ethanol precipitation. The RNA was then hydrolysed by 0.2 M NaOH at 45 °C for 60 min, the sample was neutralized and the residual nucleic acids collected by ethanol precipitation. The sample was dissolved in 10 μ l of formamide dye mix³⁶ and 3 μ l of it were loaded on to the sequencing gel. To generate molecular weight size markers, the *Bam*HI-*Hind*III end-labelled fragment was subjected to chemical sequencing reactions and products were displayed on 8% sequencing gel. Lanes A, G, C, T and AC are the sequencing ladders. Lane L, control hybridization between the probe and LTK⁻ cells total RNA; lane H, mung bean nuclease-resistant products generated by hybridization to Cd²⁺-induced HeLa cell RNA. The nucleotide sequence of the sense strand at the region of interest is indicated on the left-hand site, and the putative transcription start sites are indicated by arrows. *b*, Mapping of the termination site. ph84H DNA was digested with *Pvu*II and 3'-labelled using T₄ polymerase (BRL)⁴⁰ and [α -³²P]dCTP > 3,000 Ci mmol⁻¹ (Amersham). The labelled DNA was subjected to secondary digest with *Hind*III, and the *Pvu*II-*Hind*III fragment containing the 3' untranslated and flanking region of the MT-II_A gene was isolated as described above to be used as a probe. Hybridization, digestion and electrophoresis were as described above except for hybridization temperature which was 43 °C. End-labelled *Hpa*II fragments of pBR322 were used as size markers (lane M). Lane H, hybridization to cadmium-induced HeLa cellular RNA. Lane L, hybridization to LTK⁻ cellular RNA. Lane P, undigested probe.

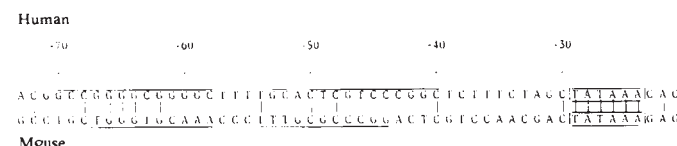


Fig. 5 Putative regulatory sequences in mammalian MT genes. The human MT-II_A and the mouse MT-I¹⁴ gene sequences are aligned, with respect to their TATAA boxes, to show the identical location of dyad symmetry sequences. TATAA sequences are boxed, and the dyad symmetry sequences are underlined. Bases are numbered from the major consensus cap site for each gene (see text). Vertical bars indicate identical bases.

it and the beginning of the pseudogene (the equivalent of the cap site of the authentic MT-II_A gene). The significance of this structure is not yet clear.

The MT-II_B pseudogene is not expressed to any significant extent in HeLa cells, and it is not expressed after introduction into mouse L cells, in experimental conditions that lead to normal expression and regulation of the MT-II_A gene (data not shown). However, such results are not surprising, because of the absence of any structures resembling eukaryotic promoters in the 5' flanking region of the MT-II_B gene.

The nucleotide sequence at the immediate 3' flanking region of the MT-II_B gene is extremely A+T-rich (90%). Such A+T-rich regions are quite prone to breakage and can conceivably generate the staggered break into which the MT-II cDNA integrated to yield the MT-II_B gene.

After re-examination of the human genomic blots shown in Fig. 1a with an MT-II_B specific probe, the polymorphism associated with the processed gene became quite apparent (Fig. 1b). In some individuals it appears as a 4.8-kb *Eco*RI fragment, the same size as the gene isolated from the genomic library, while other individuals are missing this band and instead have a 4.6-kb *Eco*RI fragment. Other individuals seem to have both of these bands, indicating that the MT-II_B pseudogene is present in two allelic forms. Of 36 individuals examined, 22 (61%) were found to be heterozygotes, 6 (17%) were homozygotes for the short form (4.6 kb) and 8 (22%) were homozygotes for the long form (4.8-kb) of the processed gene. This heterogeneity is probably due to either restriction site polymorphism around the gene or a small deletion (about 200 bp) which occurred adjacent to it.

The polymorphism is not due to rearrangement of repetitive DNA sequences near the MT-II_B, since the 4.8-kb *Eco*RI fragment bearing this gene is devoid of such sequences (M.K., unpublished results).

Several processed genes have already been described (refs 24–29 and P. Sharp, personal communication) and what seemed to be a rare event appears to be a more common phenomenon. It is interesting to note that the processed genes from the tissue-specific globin²⁸ and immunoglobulin²⁴ gene families seem to be derived from abnormal transcripts while those from the more ubiquitously transcribed families of metallothionein and tubulin (ref. 29 and P. Sharp, personal communication) are derived from apparently complete, normal transcripts. It is possible that there is a low level or occasional transcription of tissue-specific genes in the germ cells from secondary upstream or downstream promoters^{30–32}. Such aberrant transcripts are probably the precursors of the observed processed genes.

Conclusions

The human MT gene family is more complex than expected, consisting of about 11 closely related members bearing at least 80% nucleotide sequence homology to the MT-II coding region.

Not all of those genes are necessarily functional. So far we have determined the complete nucleotide sequence of two MT-II genes. While the MT-II_A gene is fully functional, the MT-II_B is a processed pseudogene.

Probably the most remarkable feature of the MT-II_A gene sequence is the lack of expected nucleotide homology to the mouse MT-I gene in the immediate 5' flanking region of both genes. Instead, the two genes reveal secondary structure homology. Both of them have dyad symmetry regions, which can form stable hairpin loops at the same positions relative to the TATAA box. So far, in the case of the mouse MT-I gene, there is evidence for the involvement of one of those structures in the metal induction phenomenon¹⁹.

This observation (secondary structure homology) might be valid for other gene systems as well. In such a case, a search for simple nucleotide homologies in the 5' flanking regions of related genes, might not reveal the important regulatory regions.

The MT-II_B pseudogene was probably generated from a normal RNA transcript of the MT-II_A gene. The structure of this and other processed genes suggests that the reversed flow

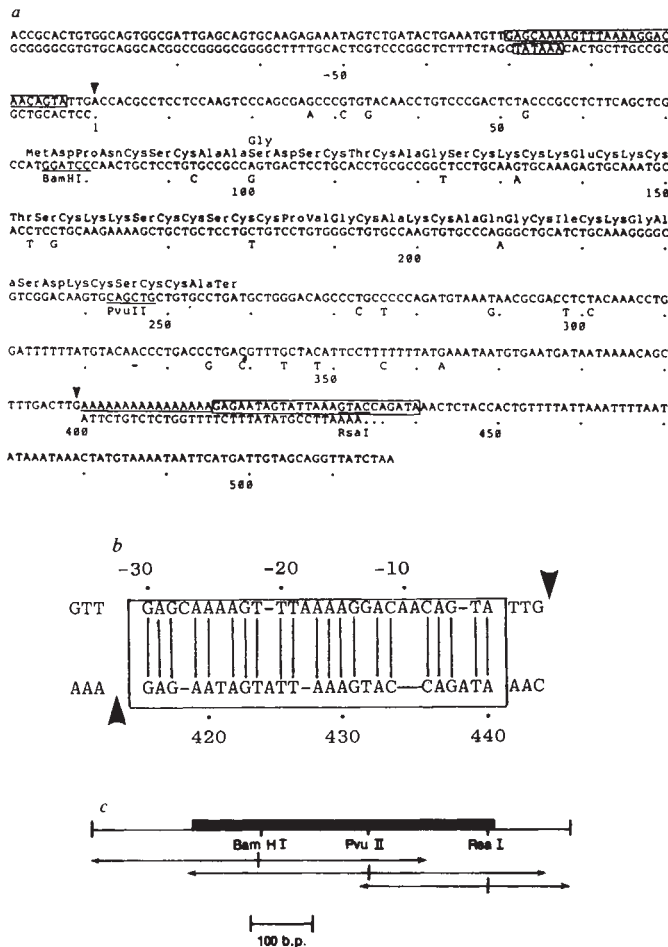


Fig. 6 Primary structure of the human MT-II_B pseudogene. *a*, The DNA sequence of the human MT-II_B pseudogene and flanking regions. Sequence analysis was performed on DNA isolated from a subclone in plasmid pBR322 (p68H) of the 4.3-kb *Hind*III restriction endonuclease fragment of ghMT-II_B. Sequence analysis was as described in Fig. 3 legend. The transcription initiation signal (–30 to –24) of the MT-II_A gene and a putative polyadenylation signal sequence are boxed. Arrows indicate the limits of homology with human MT-II mRNA. Nucleotide differences with the MT-II_A gene sequence (Fig. 3) are indicated below and amino acid differences above the pseudogene sequence. The poly(A) sequence is underlined and sequences flanking the MT-mRNA homology, containing direct repeats, are boxed. *b*, Homology of the direct repeats flanking the MT-II_B processed gene. The region of homology is boxed, identical bases are indicated by vertical bars. Arrows indicate locations of the major cap site and the termination of the poly(A) tract. *c*, Sequencing strategy for the MT-II_B processed gene. The solid box indicates the extent of the MT-II_B processed gene (including the direct repeats). Arrows indicate the direction and extent of sequencing (each sequence was determined at least twice).

of information from RNA to DNA is not limited to retroviruses, but seems to have happened on several occasions in the germ cells of mammals. Such a process might not only be responsible for the generation of non-functional processed pseudogenes, but might have accounted for the generation of functional intronless genes as well.

We thank Drs A. Guiterrez-Hartmann, C. Craik and H. E. Varmus for critical reading of the manuscript, Dr A. Gibbs for assistance with SEQ computer programs, Mr A. J. Mason and Dr. J. Shine for providing information and materials essential to the sequence analysis, Dr G. Bell for samples of human DNA and Dr J. D. Baxter for providing laboratory space. This work was supported in part by grant R-809189 from the Environmental Protection Agency (to M.K.).

Received 15 June; accepted 1 September 1982.

1. Kagi, J. H. R. & Nordberg, M. (eds) *Metallothionein* (Birkhauser, Basle, 1979).
2. Karin, M. & Herschman, H. R. *Science* **204**, 176–177 (1979).
3. Pulido, P., Kagi, J. H. R. & Vallee, B. L. *Biochemistry* **5**, 1768–1777 (1966).
4. Rudd, C. J. & Herschman, H. R. *Tox. appl. Pharmac.* **47**, 273–278 (1979).
5. Karin, M. & Herschman, H. R. *Eur. J. Biochem.* **107**, 395–401 (1980).
6. Kissling, M. M. & Kagi, J. H. R. *FEBS Lett.* **82**, 247–250 (1977).
7. Karin, M. *et al. Nature* **286**, 295–297 (1980).
8. Karin, M., Slater, E. P. & Herschman, H. R. *J. cell. Physiol.* **106**, 63–74 (1981).
9. Durnam, D. M. & Palmiter, R. D. *J. biol. Chem.* **256**, 5712–5716 (1981).
10. Hager, L. J. & Palmiter, R. D. *Nature* **291**, 340–342 (1981).
11. Karin, M. & Richards, R. *Nucleic Acids Res.* **10**, 3165–3173 (1982).
12. Lawn, R. M. *et al. Cell* **15**, 1157–1174 (1978).
13. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
14. Benton, W. D. & Davis, R. W. *Science* **196**, 180–182 (1977).
15. Glanville, N., Durnam, D. M. & Palmiter, R. D. *Nature* **292**, 267–269 (1981).
16. Breathnach, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 4853–4857 (1978).
17. Weaver, R. F. & Weissman, C. *Nucleic Acids Res.* **5**, 1175–1193 (1979).
18. Kaye, K. E., Warren, R. & Palmiter, R. D. *Cell* **29**, 99–108 (1982).
19. Brinster, R. L. *et al. Nature* **296**, 39–42 (1982).

20. Kingsbury, R. & McKnight, S. L. *Science* **217**, 316–324 (1982).
21. Larsen, A. & Weintraub, H. *Cell* **29**, 609–672 (1982).
22. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263**, 211–214 (1976).
23. Calos, M. P. & Miller, J. H. *Cell* **20**, 579–595 (1980).
24. Hollis, F. G. *et al. Nature* **296**, 321–325 (1982).
25. Leuders, K., Leder, A., Leder, P. & Kuff, E. *Nature* **295**, 426–428 (1982).
26. Van Arsdell, S. W. *et al. Cell* **26**, 11–17 (1981).
27. Jagadeeswaran, P., Forget, B. G. & Weissman, S. M. *Cell* **26**, 141–142 (1982).
28. Nishioka, Y., Leder, A. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2806–2809 (1980).
29. Wilde, C. D. *et al. Nature* **297**, 83–84 (1982).
30. Shaul, Y., Kaminichik, J. & Aviv, H. *Eur. J. Biochem.* **116**, 461–466 (1981).
31. Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937–1941 (1980).
32. Hofer, E. & Darnel, J. E. *Cell* **23**, 585–593 (1981).
33. Bell, G., Karam, J. H. & Rutter, W. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5759–5763 (1981).
34. Rigby, P. W. J. *et al. J. molec. Biol.* **113**, 237–251 (1977).
35. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).
36. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499–559 (1980).
37. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5468 (1979).
38. Goodman, H. M. *Meth. Enzym.* **65**, 63–64 (1980).
39. Heidecker, G., Messing, J. & Gronenborn, B. *Gene* **10**, 69–73 (1980).
40. O'Farrell, P. *Focus* **3**, 1–3 (1981).

LETTERS TO NATURE

A single quantum cannot be cloned

W. K. Wootters*

Center for Theoretical Physics, The University of Texas at Austin,
Austin, Texas 78712, USA

W. H. Zurek

Theoretical Astrophysics 130–33, California Institute of Technology,
Pasadena, California 91125, USA

If a photon of definite polarization encounters an excited atom, there is typically some nonvanishing probability that the atom will emit a second photon by stimulated emission. Such a photon is guaranteed to have the same polarization as the original photon. But is it possible by this or any other process to amplify a quantum state, that is, to produce several copies of a quantum system (the polarized photon in the present case) each having the same state as the original? If it were, the amplifying process could be used to ascertain the exact state of a quantum system: in the case of a photon, one could determine its polarization by first producing a beam of identically polarized copies and then measuring the Stokes parameters¹. We show here that the linearity of quantum mechanics forbids such replication and that this conclusion holds for all quantum systems.

Note that if photons could be cloned, a plausible argument could be made for the possibility of faster-than-light communication². It is well known that for certain non-separably correlated Einstein-Podolsky-Rosen pairs of photons, once an observer has made a polarization measurement (say, vertical versus horizontal) on one member of the pair, the other one, which may be far away, can be for all purposes of prediction regarded as having the same polarization³. If this second photon could be replicated and its precise polarization measured as above, it would be possible to ascertain whether, for example, the first photon had been subjected to a measurement of linear or circular polarization. In this way the first observer would be able to transmit information faster than light by encoding his message into his choice of measurement. The actual impossibility of cloning photons, shown below, thus prohibits superluminal communication by this scheme. That such a scheme must fail for some reason despite the well-established existence of long-range quantum correlations^{6–8}, is a general consequence of quantum mechanics⁹.

A perfect amplifying device would have the following effect

on an incoming photon with polarization state $|s\rangle$:

$$|A_0\rangle|s\rangle \rightarrow |A_s\rangle|ss\rangle \quad (1)$$

Here $|A_0\rangle$ is the 'ready' state of the apparatus, and $|A_s\rangle$ is its final state, which may or may not depend on the polarization of the original photon. The symbol $|ss\rangle$ refers to the state of the radiation field in which there are two photons each having the polarization $|s\rangle$. Let us suppose that such an amplification can in fact be accomplished for the vertical polarization $|\uparrow\rangle$ and for the horizontal polarization $|\leftrightarrow\rangle$. That is,

$$|A_0\rangle|\uparrow\rangle \rightarrow |A_{\text{vert}}\rangle|\uparrow\uparrow\rangle \quad (2)$$

and

$$|A_0\rangle|\leftrightarrow\rangle \rightarrow |A_{\text{hor}}\rangle|\leftrightarrow\leftrightarrow\rangle \quad (3)$$

According to quantum mechanics this transformation should be representable by a linear (in fact unitary) operator. It therefore follows that if the incoming photon has the polarization given by the linear combination $\alpha|\uparrow\rangle + \beta|\leftrightarrow\rangle$ —for example, it could be linearly polarized in a direction 45° from the vertical, so that $\alpha = \beta = 2^{-1/2}$ —the result of its interaction with the apparatus will be the superposition of equations (2) and (3):

$$|A_0\rangle(\alpha|\uparrow\rangle + \beta|\leftrightarrow\rangle) \rightarrow \alpha|A_{\text{vert}}\rangle|\uparrow\uparrow\rangle + \beta|A_{\text{hor}}\rangle|\leftrightarrow\leftrightarrow\rangle \quad (4)$$

If the apparatus states $|A_{\text{vert}}\rangle$ and $|A_{\text{hor}}\rangle$ are not identical, then the two photons emerging from the apparatus are in a mixed state of polarization. If these apparatus states are identical, then the two photons are in the pure state

$$\alpha|\uparrow\uparrow\rangle + \beta|\leftrightarrow\leftrightarrow\rangle \quad (5)$$

In neither of these cases is the final state the same as the state with two photons both having the polarization $\alpha|\uparrow\rangle + \beta|\leftrightarrow\rangle$. That state, the one which would be required if the apparatus were to be a perfect amplifier, can be written as

$$2^{-1/2}(\alpha a_{\text{vert}}^+ + \beta a_{\text{hor}}^+)^2|0\rangle = \alpha^2|\uparrow\uparrow\rangle + 2^{1/2}\alpha\beta|\uparrow\leftrightarrow\rangle + \beta^2|\leftrightarrow\leftrightarrow\rangle$$

which is a pure state different from the one obtained above by superposition [equation (5)].

Thus no apparatus exists which will amplify an arbitrary polarization. The above argument does not rule out the possibility of a device which can amplify two special polarizations, such as vertical and horizontal. Indeed, any measuring device which distinguishes between these two polarizations, a Nicol prism for example, could be used to trigger such an amplification.

The same argument can be applied to any other kind of quantum system. As in the case of photons, linearity does not forbid the amplification of any given state by a device designed especially for that state, but it does rule out the existence of a device capable of amplifying an arbitrary state.

* Present address: Department of Physics and Astronomy, Williams College, Williamstown, Massachusetts 01267, USA.

Milonni (unpublished work) has shown that the process of stimulated emission does not lead to quantum amplification, because if there is stimulated emission there must also be—with equal probability in the case of one incoming photon—spontaneous emission, and the polarization of a spontaneously emitted photon is entirely independent of the polarization of the original.

It is conceivable that a more sophisticated amplifying apparatus could get around Milonni's argument. We have therefore presented the above simple argument, based on the linearity of quantum mechanics, to show that no apparatus, however complicated, can amplify an arbitrary polarization.

We stress that the question of replicating individual photons is of practical interest. It is obviously closely related to the

quantum limits on the noise in amplifiers^{10,11}. Moreover, an experiment devised to establish the extent to which polarization of single photons can be replicated through the process of stimulated emission is under way (A. Gozzini, personal communication; and see ref. 12). The quantum mechanical prediction is quite definite; for each perfect clone there is also one randomly polarized, spontaneously emitted, photon.

We thank Alain Aspect, Carl Caves, Ron Dickman, Ted Jacobson, Peter Milonni, Marlan Scully, Pierre Meystre, Don Page and John Archibald Wheeler for enjoyable and stimulating discussions.

This work was supported in part by the NSF (PHY 78-26592 and AST 79-22012-A1). W.H.Z. acknowledges a Richard Chace Tolman Fellowship.

Received 11 August; accepted 7 September 1982.

1. Born, M. & Wolf, E. *Principles of Optics* 4th edn (Pergamon, New York, 1970).
2. Herbert, N. *Found. Phys.* (in the press).
3. Einstein, A., Podolsky, B. & Rosen, N. *Phys. Rev.* **47**, 777–780 (1935).
4. Bohm, D. *Quantum Theory*, 611–623 (Prentice-Hall, Englewood Cliffs, 1951).
5. Kocher, C. A. & Commins, E. D. *Phys. Rev. Lett.* **18**, 575–578 (1967).
6. Freedman, S. J. & Clauser, J. R. *Phys. Rev. Lett.* **28**, 938–941 (1972).

7. Fry, E. S. & Thompson, R. C. *Phys. Rev. Lett.* **37**, 465–468 (1976).
8. Aspect, A., Grangier, P. & Roger, G. *Phys. Rev. Lett.* **47**, 460–463 (1981).
9. Bussey, P. J. *Phys. Lett.* **90A**, 9–12 (1982).
10. Haus, H. A. & Mullen, J. A. *Phys. Rev.* **128**, 2407–2410 (1962).
11. Caves, C. M. *Phys. Rev. D* **15**, (in the press).
12. Gozzini, A. *Proc. Symp. on Wave-Particle Dualism* (eds Diner, S., Fargue, D., Lochak, G. & Selleri, F) (Reidel, Dordrecht, in the press).

The Crab Nebula's progenitor

Ken'ichi Nomoto*, Warren M. Sparks†, Robert A. Fesen‡, Theodore R. Gull‡, S. Miyaji‡ & D. Sugimoto*

* Department of Earth Science and Astronomy, University of Tokyo, College of General Education, 3-8-1 Komaba, Meguro, Tokyo 153, Japan

† Group X-5, Mail Stop F669, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

‡ Laboratory for Astronomy and Solar Physics, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

The study of supernovae is hampered by an insufficient knowledge of the initial stellar mass for individual supernova. Because of large uncertainties in estimating both the total mass of a remnant (including the pulsar or black hole) and any mass loss during the pre-supernova stages, the main sequence mass of the progenitor cannot be accurately determined from observations alone. To calculate an initial mass, one must rely on a combination of both theory and observation. Limits on the progenitor's mass range can be estimated by the presence of a compact remnant and comparison of the observed nebular chemical abundances with detailed evolutionary calculations¹. The Crab Nebula is an excellent choice for investigation because it contains a unique combination of characteristics: a central neutron star (pulsar) and a bright, well studied nebula having little or no swept-up interstellar material. In fact, several studies^{1–4} have suggested an initial mass of $\sim 10 M_{\odot}$ for the Crab progenitor. Recently, Davidson *et al.*⁵, quoting two of us (K.N. and W.M.S.), state that the Crab's progenitor had a mass slightly larger than $8 M_{\odot}$. Here we present in detail the reasoning behind this statement and suggest the explosion mechanism.

Briefly, the Crab consists of a pulsar (assumed here to have a mass of $\approx 1.4 M_{\odot}$) and a nebula mass of $1.2\text{--}3.0 M_{\odot}$ (refs 5, 6) which has a helium overabundance of $1.6 < X_{\text{He}}/X_{\text{H}} < 8$ (where X is an element's mass fraction). The oxygen abundance (X_{O}) is ~ 0.003 (refs 5, 6), which is less than the solar value of 0.007, while the oxygen-to-hydrogen ratio is approximately solar. The carbon-to-oxygen ratio is $0.4 < X_{\text{C}}/X_{\text{O}} < 1.1$ (ref. 5). Nitrogen may be slightly overabundant, while neon, sulphur and iron abundances are uncertain but are probably not greatly over- or underabundant. Because the Crab Nebula is helium-rich but not oxygen-rich, the hydrogen-rich (solar abundances) envelope and the helium layer of the progenitor star were ejected but the oxygen-rich layer below the helium layer was not. The lower layers must have formed the neutron star. The

lower limit of the large helium-to-hydrogen ratio means that at least half of the ejected material must have come from the helium layer.

Arnett (ref. 7 and refs therein) systematically evolved helium cores of various masses (M_{α}) into late stages of evolution. He¹ compared Davidson's⁸ derived abundances of the Crab nebula with calculated abundances from the $M_{\alpha} = 4.0 M_{\odot}$ model, which was his lowest-massed, highly evolved helium core (corresponding to approximately a $15 M_{\odot}$ star). Combining all the material above the helium-burning shell (his case B) with enough interstellar material to obtain $X_{\text{He}}/X_{\text{H}} = 8$, he found good agreement with $X_{\text{N}}/X_{\text{He}}$ and $X_{\text{O}}/X_{\text{He}}$ of Davidson's⁸ 'model 1'. However, the calculated value of $X_{\text{C}}/X_{\text{He}}$ was too large by a factor of 30. At that time, the Crab's carbon abundance had not been directly measured and Arnett suggested several possibilities: the inferred carbon abundance was too low, the carbon was hidden in the filaments, or a lower-mass helium core, $\sim 3 M_{\odot}$, was more appropriate.

Using recent UV observations with the International Ultraviolet Explorer, Davidson *et al.*⁵ have established that the carbon abundance is nearly solar. They also showed that the hydrogen and helium seemed to be fairly well mixed and, as carbon is convectively mixed in the helium layer, this would argue against carbon being hidden in the filaments. However, IR observations by Dennefeld and Andrillat⁹ showed that the strength of [C I] $\lambda 9,850$ relative to [S III] $\lambda 9,069$ varied with position in the Crab. The strongest [C I] line would indicate a rather large carbon abundance if the ionizing flux is constant. Whether the IR observations indicate variation in the carbon abundance, variation in the ionizing flux, or high densities in neutral cores is not known. For the remainder of this report we will assume the carbon abundance as determined by Davidson *et al.*⁵.

The existence of a pulsar in the Crab indicates that the progenitor's mass was larger than the upper mass limit ($8 \pm 1 M_{\odot}$)¹⁰ for degenerate carbon ignition. Degenerate carbon ignition results in carbon deflagration¹¹ which completely disrupts the star, leaving no compact remnant. Lower-mass stars that lose enough mass to avoid degenerate carbon burning eventually become white dwarfs. Stars massive enough ($\geq 8 M_{\odot}$) to burn carbon non-degenerately will eventually undergo a core collapse initiated either by electron capture¹² onto Mg, Ne and O or by burn-out of all the available fuel^{13,14}. When the collapsing core reaches neutron-star densities, stability is regained. Although detailed calculations of the collapse remain inconclusive, it is generally felt that the core will overshoot its equilibrium position and then rebound, initiating a shock wave⁴. This shock wave ejects the outer material but not the core, resulting in both a supernova nebula and a pulsar¹⁵. In more massive stars

the core will continue to collapse down to a black hole, but the stellar mass at which this starts to happen is unknown⁴. Therefore, the presence of a pulsar in the Crab establishes a lower mass limit of $8 \pm 1 M_{\odot}$, but cannot set an upper limit for the progenitor.

We will now investigate the mass constraints due to the chemical abundances of the nebula. Stars more massive than $\sim 10 M_{\odot}$ will undergo an iron core collapse while less massive stars (but more massive than the carbon deflagration supernova) will undergo an electron capture core contraction¹². Let us consider the lowest mass star which undergoes an iron core collapse, that is, $10 M_{\odot}$. Recent evolutionary calculations³ of such a star show that in the late stages there is a $1.5 M_{\odot}$ core, which presumably forms the neutron star, a helium shell of $1.2 M_{\odot}$ and a $7.3 M_{\odot}$ hydrogen-rich envelope. If at least $6 M_{\odot}$ of the hydrogen-rich envelope were lost in pre-supernova stages, the supernova would eject a helium-rich nebula like that of the Crab. In fact, Woosley *et al.*³ and Hillebrandt⁴ suggested this model to fit the Crab. However, the recently observed carbon abundances⁵ also contradict this model. During the late stages of evolution, the helium burning shell becomes so active that it forms a convective region and deposits appreciable amounts of carbon (and oxygen to a less extent) in the helium layer. In this model the convective region included $0.7 M_{\odot}$ of the $1.2 M_{\odot}$ helium layer and contained 0.04 carbon. Thus, if the helium layer were mixed with the maximum allowable amount of hydrogen-rich material ($1.2 M_{\odot}$), the mixture would contain $X_c \sim 0.012$, which is larger than that observed. For somewhat larger-mass stars the mass considerations are similar but the amount of carbon in the helium layer increases even further.

Recently, evolutionary models of the helium cores of stars from 8 to $10 M_{\odot}$ have become available. In addition to the published evolution of helium core masses of 2.6 and $2.8 M_{\odot}$ (ref. 16), one of us (K.N.) has evolved a helium core mass of $2.4 M_{\odot}$. (In these models, the stellar mass is about a factor of 4 times the helium core mass, which is then allowed to increase to simulate the core growth due to the hydrogen-burning shell, and the core mass values quoted are at the end of the central helium burning.) Figure 1a shows the chemical evolution for the $2.6 M_{\odot}$ helium core. During the late stages, a large convective region develops above the helium burning shell and deposits a $0.02 M_{\odot}$ of carbon in the helium layer. This is more carbon than the observations allow and a lower-mass helium core is necessary. Another interesting phenomenon occurs slightly later than the development of the helium-burning convective region. As the carbon-burning shell approaches the helium-burning shell, the helium layer expands greatly. This allows the surface convective region to penetrate the helium layer and dredge most of it up.

The dredge-up phase has a more significant role in the $2.4 M_{\odot}$ helium core model, as demonstrated in Fig. 1b. (In this computation, the hydrogen-rich envelope of $1.1 M_{\odot}$ was fitted to the outer edge of the $2.4 M_{\odot}$ helium core; that is, the stellar mass was assumed to have decreased from the initial mass of $\sim 9 M_{\odot}$ to $3.5 M_{\odot}$ by mass loss.) Because of the stronger electron-degeneracy, the penetration of the surface convective region starts earlier than in the $2.6 M_{\odot}$ helium core case. (A similar penetration takes place for a $9 M_{\odot}$ star¹⁷.) Afterwards, the dredge-up of the helium layer proceeds as follows: (1) the first dredge-up reduces the mass of the helium layer from $1.1 M_{\odot}$ to $0.14 M_{\odot}$, (2) a temporal retreat of the surface convection zone during the several carbon shell flashes in the core, (3) a second dredge-up which reduces the helium layer mass to $5 \times 10^{-3} M_{\odot}$, and (4) retreat of the surface convection zone with re-ignition of the hydrogen burning shell. Further evolution of the core mass growth from $1.30 M_{\odot}$ to $1.38 M_{\odot}$ with the hydrogen-helium double shell burnings will be similar to the growth of the carbon-oxygen cores¹⁰.

Between the dredge-up phases, a relatively small helium shell flash occurs and produces $\sim 0.007 M_{\odot}$ carbon. However, the strength of the shell flash has a positive correlation with stellar

mass. If the first phase dredges up more helium, the amount of carbon produced would be smaller. In fact, the amount of carbon produced by the $2.4 M_{\odot}$ helium core model is one-third of that produced by the $2.6 M_{\odot}$ helium core model. Therefore, the carbon abundance not only decreases with decreasing stellar mass but also the abundance has a rather sharp drop-off caused by the presence or absence of the first dredge-up phase. The Crab's observed carbon abundance is consistent with its progenitor's mass being below the critical mass where this sharp drop-off occurs. The exact value of the critical stellar mass at which this occurs must await more detailed evolutionary models, but we estimate it to be $9.5 \pm 0.5 M_{\odot}$.

We therefore believe that the mass of the Crab's progenitor is between the upper limit for carbon deflagration ($8 \pm 1 M_{\odot}$) and the lower limit set by the dredge-up of the helium layer before development of the helium-burning convective region ($9.5 \pm 0.5 M_{\odot}$). According to the latest evolutionary models this mass range falls entirely within the stars that undergo electron capture supernova¹². This would make the Crab the remnant of an electron capture supernova.

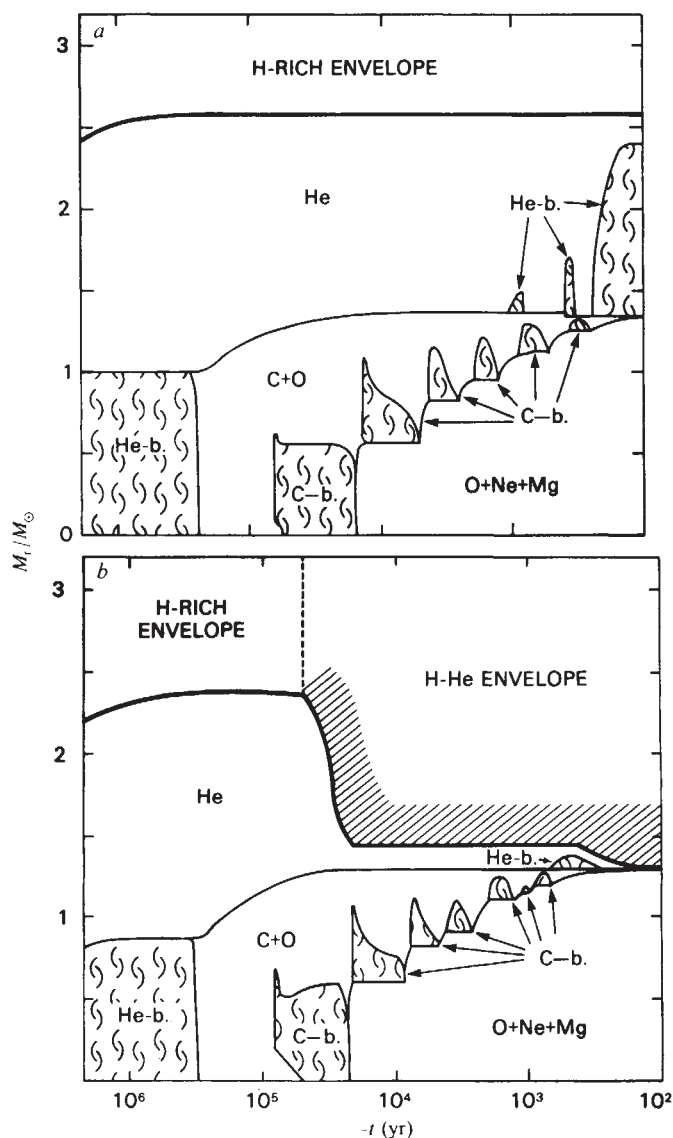


Fig. 1 a, Chemical evolution of the star with a $2.6 M_{\odot}$ helium core. The shaded regions are convective. Time, t , is measured from the onset of the penetration of the surface convection zone into the helium layer. b, Same as a, but for the star with a $2.4 M_{\odot}$ helium core. Time, t , is measured from the stage when the dredge up of the helium layer stops and the hydrogen shell burning is re-ignited.

Based on the above discussion, we propose the following scenario for the Crab's progenitor star as being plausible and most consistent with observations. On the main sequence the star had an initial mass of between 8 and $9.5 M_{\odot}$, and spent 3×10^7 yr, 2×10^6 yr and 6×10^4 yr in hydrogen, helium and carbon burning stages, respectively. During the blue and red supergiant phases, the star lost $5-6 M_{\odot}$ by a stellar wind, reducing the hydrogen-rich envelope to $0.5-1.0 M_{\odot}$. During carbon burning the helium layer expanded and allowed the surface convective region to penetrate the helium layer. Most of the helium layer ($\sim 1.2 M_{\odot}$) was dredged up into the hydrogen-rich layer, mixing their compositions. As the degenerate 0-Ne-Mg core approached $1.38 M_{\odot}$, electron captures onto Mg, Ne and O reduced the Chandrasekhar's limiting mass to below the core mass, causing a rapid contraction. During the contraction, the remainder of the nuclear fuel in the core burned, and practically all of the energy released escaped as neutrinos. The contracting core finally became free-fall and when it reached neutron-star densities, the core overshot its equilibrium position and then rebounded, creating a shock wave. This shock wave was not strong enough to eject the core, but it was strengthened by the steep density gradient at the core edge enough to eject the loosely bound hydrogen-helium envelope. The $1.38 M_{\odot}$ core was not massive enough to collapse down to a black hole but remained a neutron star, in agreement with the observations. There was little or no nuclear burning in the hydrogen-helium envelope due to the passage of the shock wave. However, the propagation of the shock wave through the extended hydrogen-helium envelope caused the rapid rise and fall in the supernova light curve¹⁸. The ejected hydrogen-helium layer had a mass of $\sim 2 M_{\odot}$, a composition of $X_H \sim 0.2-0.3$, $X_{He} \sim 0.8-0.7$, its original solar ratio of X_O/X_H , but slightly enhanced carbon as $X_C/X_O \sim 1$. Because the helium layer was processed via the CNO cycle, the X_N/X_H ratio was ~ 10 times solar. These abundances are in agreement with many of the observations of the Crab nebula. The suggested abundance variations^{6,19} among the filaments could be due to one or more of the following: (1) incomplete mixing during dredge-up, (2) helium shell flashes in the thin, residual layer of helium after the first dredge-up, or (3) mixing with previously ejected hydrogen-rich circumstellar material. It is possible that the rebounding shock strength in the electron-capture supernovae will be weaker than the iron core collapse supernovae because the electron capture core slowly approaches free-fall while the iron core quickly reaches free-fall. Thus, the relatively low kinetic energy of the Crab Nebula may also be consistent.

To establish the correctness of the above scenario, we need actual modelling that shows that the rebounding shock ejects the envelope and gives rise to the supernova light curve. Although most of the recent calculations for the collapse of massive stars ($>15 M_{\odot}$) fail to obtain an explosion¹⁵, the recent calculation of a $10 M_{\odot}$ star explosion by Hillebrandt⁴ is very suggestive. Because a $10 M_{\odot}$ star has a smaller iron core mass than a $15 M_{\odot}$ star, the outgoing shock wave is strengthened by the steep density gradient near the silicon burning shell (or the outer edge of the iron core) and ejects a relatively small amount of core material. In $8-10 M_{\odot}$ stars, much steeper density gradients exist at the helium burning shell, making it more likely that the shock is strengthened there and thus ejecting only a helium-rich layer.

We cannot be certain whether the mechanism described above applies to other supernova remnants. Although a few objects are considered to be 'Crab-like' remnants (for example, 3C58), this classification has been based largely on their radio emission properties. Until we know more about these remnants (their elemental abundances, total kinetic energies and whether they contain neutron stars), we cannot determine how similar they are to the Crab Nebula. Therefore, although the Crab Nebula was probably the product of an electron capture supernova, it is not yet clear if other Crab-like supernova remnants originated by a similar explosion mechanism.

Received 6 May; accepted 3 August 1982.

- Arnett, W. D. *Astrophys. J.* **195**, 727-733 (1975).
- Gott, J. R., Gunn, J. E. & Ostriker, J. P. *Astrophys. J. Lett.* **160**, L91-L96 (1970).
- Woosley, S. E., Weaver, T. A. & Taam, R. E. in *Type I Supernovae* (ed. Wheeler, J. C.) 96-112 (University of Texas, Austin, 1980).
- Hillebrandt, W. in *Supernovae* (eds Rees, M. J. & Stoneham, R. J.) 123-128 (Reidel, Dordrecht, 1982).
- Davidson, K. *et al. Astrophys. J.* **253**, 696-706 (1982).
- Henry, R. C. & MacAlpine, G. M. *Astrophys. J.* **258**, 11-21 (1982).
- Arnett, W. D. *Astrophys. J. Suppl.* **35**, 145-160 (1977).
- Davidson, K. *Astrophys. J.* **186**, 223-231 (1973).
- Dennefeld, M. & Andriolat, Y. *Astr. Astrophys.* **103**, 44-49 (1981).
- Sugimoto, D. & Nomoto, K. *Space Sci. Rev.* **25**, 155-227 (1980).
- Nomoto, K., Sugimoto, D. & Neo, S. *Astrophys. Space Sci.* **39**, L37-L42 (1976).
- Miyaji, S., Nomoto, K., Yokoi, K. & Sugimoto, D. *Publ. astr. Soc. Japan* **32**, 303-329 (1980).
- Hoyle, F. & Fowler, W. A. *Astrophys. J.* **132**, 565-590 (1960).
- Colgate, S. A. & White, R. H. *Astrophys. J.* **143**, 626-681 (1966).
- Hillebrandt, W. *Astr. Astrophys.* **110**, L3-L6 (1982).
- Nomoto, K. in *Fundamental Problems in the Theory of Stellar Evolution* (eds Sugimoto, D., Lamb, D. Q. & Schramm, D. N.) 295-315 (Reidel, Dordrecht, 1981).
- Weaver, T. A., Axelrod, T. S. & Woosley, S. E. in *Type I Supernovae* (ed. Wheeler, J. C.) 113-154 (University of Texas, Austin, 1980).
- Falk, S. W. & Arnett, W. D. *Astrophys. J. Suppl.* **33**, 515-562 (1977).
- Fesen, R. A. & Kirshner, R. P. *Astrophys. J.* **258**, 1-10 (1982).

Siderophiles in the Brachina meteorite: impact melting?

Graham Ryder

Northrop Services, Inc., Lunar Curatorial Laboratory,
PO Box 34416, Houston, Texas 77034, USA

The published high abundances and unusual ratios of siderophiles in the chassignite Brachina (refs 1, 2 and R. A. Schmitt, personal communication in ref. 1) in comparison with its apparent relatives Chassigny, shergottites, and nakhlites²⁻¹¹, suggest that Brachina might be an impact melt. If so, the projectile was chemically similar to either some iron meteorites¹² or the Eagle Station trio¹³ of pallasites. As contaminants, the latter meteorites would also explain the deviant oxygen isotope ratios¹⁴ of Brachina without recourse to a separate parent planet. We show here that all these meteorites could have formed on Mars, with Brachina ejected as a large tektite-like melt blob, along with shocked and unshocked target rocks. The hypothesis creates predictions which can be tested with data from radiometric ages and dynamic crystallization experiments.

Shergottites, nakhlites and chassignites are distinct from other achondrites, particularly in their higher oxidation states, higher volatile abundances, and young crystallization ages of 1,300 Myr. These meteorites are interesting because of their possible origin as igneous rocks on Mars¹⁵⁻¹⁹. One characteristic apparently not consistent with this is that the ratios of oxygen isotopes of Brachina are distinct from those of all the other shergottites, nakhlites and Chassigny (which fall within error of a mass fractionation line) (Fig. 1), and requires a separate parent planet for Brachina¹⁴. One other feature of Brachina is its texture, which is dominated by fine-grained ($\sim 200 \mu\text{m}$) olivines, which are, however, not quench-textured^{20,21}. The texture is dissimilar to known igneous rocks.

An additional peculiar characteristic of Brachina appears to have passed unnoticed: its extremely high siderophile abundances and their peculiar ratios, unlike shergottites, nakhlites, and Chassigny, or lunar and terrestrial basalts (Table 1). All the abundances for Brachina are high; the most conspicuously different ratios are the high Ni/Co and Ir/Au. The Ni/Co ratio is much closer to chondritic, stony-irons, and irons than to other differentiated silicate materials, for example. Among meteorite groups, the Ir/Au ratio is distinctive (Fig. 2), and the two analyses of Ir and Au for Brachina are remarkably consistent. (However, data for Shergotty and Nakhla are quite varied and inconsistent; Table 1.) These siderophile data, like the oxygen isotopes, demonstrate that Brachina is different in origin from the other shergottites, nakhlites, and chassignites.

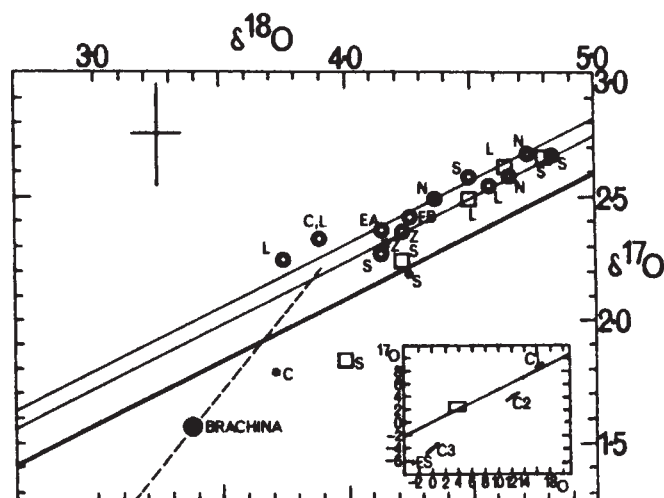


Fig. 1 Oxygen isotope data for Brachina, Chassigny, shergottites, and nakhlites. Heavy diagonal line is terrestrial mass fractionation line; light diagonal lines are reasonable mass fractionation lines for Shergottites, nakhlites, and Chassigny. Dashed line through Brachina connects with Eagle Station trio pallasites. Large cross is precision²³. S, Shergotty; Z, Zagami; L, Lafayette; N, Nakhla; EA, EB, EETA79001; C, Chassigny. Inset shows main diagram area as a rectangle, with Eagle Station trio and carbonaceous chondrite groups for comparison. Data references: stars¹⁴; squares²⁵; asterisks, Clayton (personal communication in ref. 17); Eagle Station and chondrites²³.

One possible explanation of the siderophile data is that Brachina is from the same planet as Chassigny, shergottites, and nakhlites, but that it is an impact melt, contaminated with other siderophile-bearing meteoritic material. For such an hypothesis, by assuming indigenous abundances for Brachina similar to those of Chassigny, one can estimate the abundances resulting from contamination (Table 2); the Ni/Co ratio is similar to most meteorite groups except achondrites. For reasonable amounts of contamination (there is $\leq 10\%$ in the most-siderophile-enriched terrestrial impact melts), stony

meteorites are eliminated because they would require more than 25% contamination and their Ir/Au ratios are too low. Of the remaining meteorites, a few of the iron meteorite groups¹² and the Eagle Station trio pallasites¹³ (Table 2) can reasonably account for the siderophiles. Mesosiderites might account for the siderophiles, but Au data are not available. The Eagle Station trio of pallasites actually have Ni/Ir only half that of the hypothesized contaminant, but the ratio for the contaminant is not precise. Differences in the abundances of sulphide phases between analysed splits of Brachina could account for variations in absolute abundances and ratios in the analyses. The available data have varied nickel concentrations, and suggest that within Brachina, Ni and Ir are not concentrated in a single phase.

Of the contaminant possibilities, the Eagle Station trio of pallasites has unusual oxygen isotope characteristics, and Brachina falls on a mixing line between the Eagle Station trio and a point near Chassigny (Fig. 1). The hypothesized contaminant requires $\sim 2\%$ of Eagle Station trio metal to account for the Ni and Co excess in Brachina, or 1% to account for the Ir and Au. The Eagle Station trio are unusual in having a silicate/metal ratio of 2:1 (ref. 22), thus providing a corresponding 2–4% contamination with silicates. The oxygen isotope data (Fig. 1) require a silicate contamination of 4–11% (from $\delta^{17}\text{O}$) or 5–10% (from $\delta^{18}\text{O}$), according to the isotopic precisions²³, and with a reasonable assumption of the oxygen isotopic composition of the indigenous material. Thus the mixing of about 6 or 7% Eagle Station trio-like material and Chassigny-like material would provide a composition with oxygen isotopes and siderophile abundances and ratios similar to, though not precisely those of, Brachina. The fit would be more exact if (1) Eagle Station trio pallasites actually have a higher silicate/metal ratio than 2:1 (the slab of Eagle Station depicted in ref. 22 appears to have a higher proportion of silicates); or (2) the oxygen isotopic composition of Brachina actually lies slightly closer to the shergottite–nakhlite–chassignite mass fractionation line, which is possible; some other data vary outside of analytical precision, and a duplicate measurement for Brachina is desirable; or (3) Brachina actually has higher abundances of Ir and Au. Contamination with iron meteorites, main

Table 1 Published siderophile element abundances for shergottites, nakhlites, chassignites, lunar basalts and terrestrial basalts

Sample	Ref.	Ni (p.p.m.)	Co (p.p.m.)	Ir (p.p.b.)	Au (p.p.b.)	Ni/Co	Ni/Ir ($\times 10^3$)	Ni/Au ($\times 10^3$)	Ir/Au
Shergotty	2	83	39		16	2.1	(47)	5.2	
	3		63	1.55	88	(1.2)		(0.82)	0.02
	4	56	35			1.6			
	6		62			(1.2)			
	7	80		0.072	0.98	(1.6)	1,111	81	0.07
Zagami	3			0.10	2.10		(670)	(32)	0.05
	4	67	37			1.8			
ALHA 77005	5	320	70			4.6			
Nakhla	2	90	54	<2	2.9	1.7	>45	31	<0.68
	3		81	17	0.55	(1.1)	(5.3)	(163)	31
				718	2.5		(<11)	(36)	>3.2
	6		45			(2)			
Lafayette	3			0.13	21.0				0.01
	6		44						
Chassigny	2	480	126	2.4	0.9	3.8	200	533	2.7
	9	400	124	6	6	3.2	67	67	1
	10	475	141			3.4			
	11	~600				(~4.6)			
Brachina	2	1,200	234	130	10	5.1	9.2	120	13
	2	4,180	265	111	15	15.8	38	278	7.4
	1	3,600	320			11.3			
Lunar low-Ti basalts	26	40	40	0.01	0.03	1	4,000	1,500	0.33
Terrestrial basalts	27	120	40	0.02	0.02	3	6,300	500	0.08

Values in parentheses use averages of other splits for one parameter. p.p.b., Parts per 10^9 .

Table 2 Composition of hypothetical Brachina contaminant and Eagle Station trio pallasites¹³

	Brachina contaminant	Eagle Station	Itzawisis	Cold Bay
Ni (p.p.m.)	2,500	154,000	147,000	136,000
Co(p.p.m.)	140	8,000	8,100	5,400
Ir (p.p.b.)	115	11,400	16,000	5,800
Au(p.p.b.)	10	980	1,040	870
Ni/Co	18	19.3	18.1	25.2
Ni/Ir ($\times 10^3$)	22	13	9	23
Ni/Au ($\times 10^3$)	250	157	141	156
Ir/Au	11.5	11.6	15.4	6.7

group pallasites, or mesosiderites would have a negligible effect on the oxygen isotopes.

An alternative to the impact melt explanation for the siderophile data is that Brachina is an igneous rock containing cumulate sulphides, bearing Ni, Co, Ir and Au as chalcophile elements. Such an origin could possibly explain generally high siderophile abundances, but probably not explain the Ir/Au ratios. Au and Ir are similarly chalcophile²⁴, and thus the liquid from which Brachina accumulated would be required to have Ir/Au similar to Brachina, hence still unlike shergottites, nakhlites, and Chassigny. Without detailed knowledge of the comparative chalcophilic tendencies of Ir and Au, such an origin cannot be discarded, but does not explain the oxygen isotope data as does the contamination hypothesis.

Brachina (unlike Chassigny) is unshocked, a characteristic expected of an impact melt. Its unusual texture is not unlike some chondrules, for example porphyritic and granular ones. The inclusions^{20,21} in the cores of olivines and plagioclases are not unexpected in impact melts; the effect of these inclusions on nucleation could strongly influence texture. Unfortunately, there are no data from dynamic experiments on appropriate compositions crystallizing with heterogeneous nucleation. Brachina could have formed in a blob of melt ejected from the parent planet in the event which ejected the shergottites, nakhlites, and Chassigny ~180 Myr ago; the blob would be analogous to a large, ejected tektite. If the parent planet is Mars, then it is quite possible that the high S abundance of Brachina is derived from the martian regolith. While Brachina does not have the particularly high chlorine which might be

expected, and the presumed sulphate of the regolith would have to be reduced to the sulphide of Brachina, substantial devolatilization and reduction is normal in the production of terrestrial tektites.

If the hypothesis is correct, Brachina will have a crystallization age of ~180 Myr, apparent in internal isochrons and Ar-Ar ages. There should be no sign of the 1,300 Myr event (shown by its relatives) in the Ar data, but Sm-Nd data should be consistent with a 1,300 Myr event. Rb-Sr whole rock data might be affected by volatilization of Rb during the melting.

The hypothesis that Brachina is a melted mixture of a several per cent Eagle Station trio-like pallasite material and Chassigny-like target material reasonably accounts for both the siderophile and oxygen isotopic data. The challenge for those postulating a purely igneous origin is to explain the high siderophiles and their peculiar ratios, even if one invokes a parent planet separate from that of shergottites, nakhlites and Chassigny to account for the oxygen isotopes. The hypothesis of an origin as an impact melt blob ejected from Mars along with shergottites, nakhlites, and Chassigny is consistent with known data, and can be tested with attainable data, such as Ar-Ar ages, more complete siderophile data for all these meteorites, and experimental reproduction of the textures.

This work was supported by NASA 15425 for the operation of the Lunar Curatorial Laboratory. I thank C. A. Wood and J. Gooding for helpful comments.

Received 20 July; accepted 6 September 1982.

1. Johnson, J. E. *et al. Rec. South Austral. Mus.* **17**, 309-319 (1977).
2. Dreibus, G. *et al. Lunar planet. Sci.* **13**, 186-187 (1982).
3. Laul, J. C. *et al. Geochim. cosmochim. Acta* **36**, 329-345 (1972).
4. Stolper, E. & McSween, H. Y. Jr *Geochim. cosmochim. Acta* **43**, 1475-1498 (1979).
5. Ma, M.-S. *et al. Proc. Lunar planet. Sci.* **12B**, 1349-1358 (1981).
6. Schmitt, R. A. *et al. Meteoritics* **7**, 131-214 (1972).
7. Chou, C.-L. *et al. Proc. 7th Lunar Sci. Conf.* 3501-3518 (1976).
8. Ehmann, W. D. *et al. Geochim. cosmochim. Acta* **34**, 493-507 (1970).
9. Boynton, W. V. *et al. Geochim. cosmochim. Acta* **40**, 1439-1447 (1976).
10. Jérôme, D. Y. thesis, Univ. Oregon (1970).
11. Dyakonova, M. I. & Kharitonova, V. Y. *Meteoritika* **18**, 48-67 (1960).
12. Scott, E. R. D. *Earth planet. Sci. Lett.* **37**, 273-284 (1977).
13. Scott, E. R. D. *Geochim. cosmochim. Acta* **41**, 349-360 (1977).
14. Clayton, R. N. & Mayeda, T. K. *Lunar planet. Sci.* **13**, 117-118 (1982).
15. Wasson, J. T. & Wetherill, G. W. in *Asteroids* 926-974 (University of Arizona Press, Tucson, 1979).
16. McSween, H. Y. Jr & Stolper, E. *Scient. Am.* **242**(6), 54-63 (1980).
17. Wood, C. A. & Ashwal, L. D. *Proc. Lunar planet. Sci.* **12B**, 1359-1375 (1981).
18. Ashwal, L. D., Warner, J. L. & Wood, C. A. *Proc. Lunar planet. Sci.* **13** (in the press).
19. Nyquist, L. *Proc. Lunar planet. Sci.* **13** (in the press).
20. Nehru, E. C. *et al. Meteoritics* **14**, 493-494 (1979).
21. Floran, R. J. *et al. Geochim. cosmochim. Acta* **42**, 1213-1229 (1978).
22. Mason, B. *Meteorites* (Wiley, New York, 1962).
23. Clayton, R. N. & Mayeda, T. K. *Earth planet. Sci. Lett.* **40**, 168-174 (1978).
24. Greenland, L. P. *Geochim. cosmochim. Acta* **35**, 319-322 (1971).
25. Clayton, R. N. *et al. Earth planet. Sci. Lett.* **30**, 10-18 (1976).
26. Delano, J. W. & Ringwood, A. E. *Proc. 9th Lunar planet. Sci. Conf.* 111-159 (1978).
27. Wolf, R. & Anders, E. *Geochim. cosmochim. Acta* **44**, 2111-2124 (1980).

Terrestrial-type xenon in meteoritic troilite

Golden Hwaung & Oliver K. Manuel

Department of Chemistry, University of Missouri, Rolla, Missouri 65401, USA

Recently it has been realized that the isotopic composition of Xe in different phases of chondrites is not uniform and that AVCC Xe is a mixture of the different nucleogenetic components trapped in these phases¹⁻⁴. We show here a similar abundance pattern for the nonradiogenic xenon isotopes in air and in meteoritic troilite (FeS), which suggests that the isotopic composition of atmospheric Xe was not produced by unique events in the history of terrestrial material but represents a particular mix of the different nucleogenetic components of Xe that was dominant in a central Fe- and S-rich region of the protoplanetary nebula⁵.

Radiogenic ¹²⁹Xe has been observed in samples from the Earth's mantle, but the abundance pattern of the eight

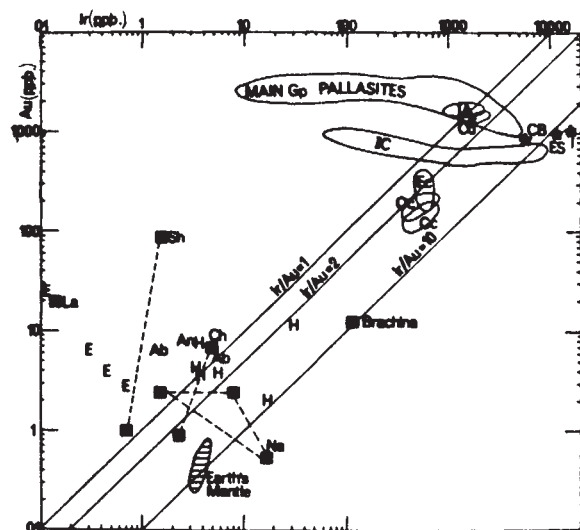


Fig. 2 Au/Ir for several meteorites. Black squares are shergottites, nakhlites, and chassignites (data in Table 1): Sh, Shergotty; Ch, Chassigny; Na, Nakhla; La, Lafayette. Stars are Eagle Station trio pallasites (data from ref. 13): ES, Eagle Station; I, Itzawisis; CB, Cold Bay. Other meteorites (data from many sources): E, Euclites; H, Howardites; Ab, Aubrites; An, Angrite; IA, IC, irons; Cd, Canyon Diablo; Oc, ordinary chondrites; Ec, enstatite chondrites; Cc, Carbonaceous chondrites. p.p.b., parts per 10⁹.

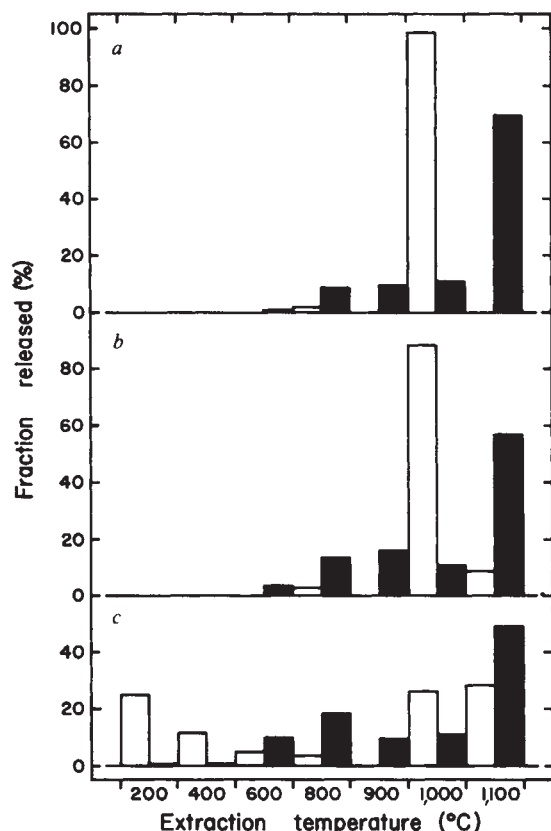


Fig. 1 The per cent of ^{132}Xe (c) and the neutron-capture-produced ^{131}Xe (b) and ^{83}Kr (a) released at each extraction temperature from troilite of the Odessa (open bars) and Cape York (black bars) iron meteorites. Note that Cape York is shown on the right and Odessa on the left for each extraction temperature, although the extraction schedules were identical except for the additional 900 °C step used only for Cape York. Values for ^{131}Xe and ^{83}Kr are defined by equations (1) and (2), respectively. Extraction temperatures of 200, 400, 600, 800, 1,000 and 1,100 °C were used for both Xe and Kr samples, and an additional gas fraction at 900 °C was collected and analysed from the Cape York sample. The concentrations of ^{132}Xe shown were measured by the peak height method and have uncertainties of $\pm 6\%$. Uncertainties of ± 50 °C are estimated for extraction temperatures above 700 °C.

nonradiogenic isotopes is atmospheric rather than chondritic⁶. 'Terrestrial' Xe will therefore be used instead of 'atmospheric Xe' in referring to the relative abundances of the nonradiogenic xenon isotopes of this planet.

Contamination has been suggested^{4,7} as an explanation for previous observations of terrestrial Xe in meteoritic and lunar material^{4,8-15}. In several cases a spike of terrestrial Xe was observed in the gas released at ≈ 900 –1,100 °C when sulphide separates of chondrites and iron meteorites were subjected to stepwise gas extraction^{4,8,9}. Samples of meteoritic troilite (FeS) were selected for this investigation because (1) this is an abundant sulphide mineral of meteorites and large nodules of troilite are easily recovered from iron meteorites, (2) adsorbed atmospheric gases can be selectively released in the low-temperature fractions when troilite is heated under vacuum, and (3) troilite nodules contain traces of selenium and tellurium, and (n, $\gamma\beta$ -) reactions on ^{82}Se and ^{130}Te have produced excess ^{83}Kr and ^{131}Xe within the troilite during the exposure of the meteoroid to cosmic rays. These two stable end products of neutron-capture, ^{83}Kr and ^{131}Xe , can be used as tracer isotopes to identify the temperature at which gases are released from within the troilite.

For this study, we separated nodules of troilite from two large meteorites, the Odessa Texas coarse octahedrite and the Cape York, West Greenland medium octahedrite. Procedures described elsewhere¹⁶ were used to analyse the isotopic composition of Kr and Xe in each gas fraction as a sample was heated to successively higher temperatures.

Experimental details and the isotopic composition of Xe in each temperature fraction are shown in Table 1 together with those of atmospheric¹⁷ and AVCC¹⁸ xenon. The concentrations and isotopic composition of Kr are given elsewhere¹⁹. Spallation and neutron-capture reactions can account for all observed deviations from the isotopic composition of atmospheric Kr.

To identify the gas component released from within the troilite, the amount of excess ^{131}Xe from the $^{130}\text{Te}(n, \gamma\beta^-)$ ^{131}Xe reaction and the amount of excess ^{83}Kr from the $^{82}\text{Kr}(n, \gamma\beta^-)$ ^{83}Kr reaction were computed as follows:

$$^{131}\text{Xe} = [^{132}\text{Xe}][(^{131}\text{Xe}/^{132}\text{Xe})_{\text{Fes}} - (^{131}\text{Xe}/^{132}\text{Xe})_{\text{Air}}] \quad (1)$$

$$^{83}\text{Kr} = [^{86}\text{Kr}][(^{83}\text{Kr}/^{86}\text{Kr})_{\text{Fes}} - (^{83}\text{Kr}/^{86}\text{Kr})_{\text{Air}}] \quad (2)$$

The release patterns of ^{132}Xe and neutron-capture-produced ^{131}Xe and ^{83}Kr for the two troilite samples are shown in Fig. 1.

The total concentration of xenon released from the Cape York troilite, $^{132}\text{Xe} = 1.79 \times 10^{-12} \text{ cm}^3 \text{ STP g}^{-1}$, was appreciably less than that released from the Odessa troilite, $11.5 \times 10^{-12} \text{ cm}^3 \text{ STP g}^{-1}$. Only a small fraction of this gas (1.4%) was released from Cape York troilite below 600 °C, whereas 36.7% of the ^{132}Xe was released from Odessa troilite below 600 °C. The release of significant amounts of xenon from Odessa troilite at low temperatures and the high concentration of ^{132}Xe in this sample seem to be compatible with the earlier finding that Odessa is an agglomeration of heterogeneous material that has not been molten since accretion whereas Cape York had an igneous origin and appears to have undergone an extended period of annealing^{20,21}.

From the release patterns of ^{83}Kr and ^{131}Xe , we conclude that degassing of troilite is represented by the 1,000 °C fraction of Odessa and the 1,100 °C fraction of Cape York. Figure 1c shows that an appreciable fraction of the total Xe was released when the troilite melted. The amount of Xe in the 1,000 °C fraction of Odessa and the 1,100 °C fraction of Cape York exceeds by more than a factor of 50 the amount of Xe observed in blanks, when the empty crucible was heated to 1,100 °C and the gases collected, cleaned and analysed by the same procedure used for the samples. Adsorbed atmospheric Xe would not be expected to survive the lower extraction temperatures, and the abundance pattern of the nonradiogenic Xe isotopes in the 1,000 °C fraction of Odessa troilite and the 1,100 °C fraction of Cape York troilite can therefore be used to determine the isotopic composition of Xe trapped in these samples.

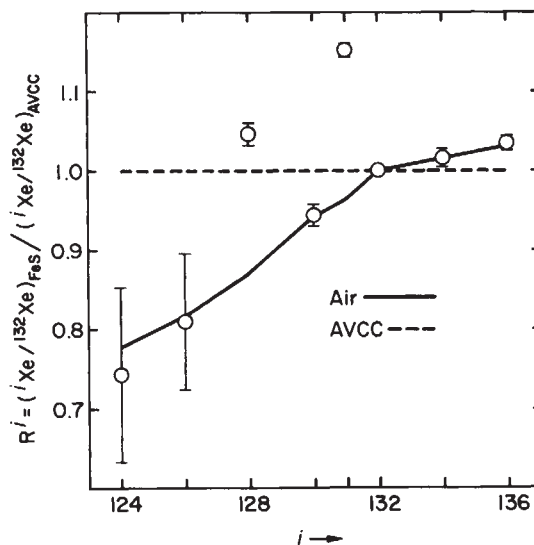


Fig. 2 A comparison of the isotopic composition of Xe in the 1,000 °C fraction of Odessa troilite (O) with that in air and in chondritic meteorites¹⁸, AVCC. AVCC Xe is represented by the dashed line at $R = 1.00$. The isotopic composition of Xe released in this melt fraction of Odessa troilite is identical to that in air, except for excess ^{128}Xe and ^{131}Xe in the troilite. These are the stable end products of neutron-capture on ^{127}I and ^{130}Te .

Table 1 Isotopic composition of xenon and concentration of ^{132}Xe released by stepwise heating of meteoritic troilite

Extraction temperature (°C)	Isotopic composition, ($^i\text{Xe}/^{132}\text{Xe}$) $\times 10^2$								^{132}Xe in units of $10^{-12}\text{ cm}^3\text{ STP g}^{-1}$
	$i = 124$	$i = 126$	$i = 128$	$i = 129$	$i = 130$	$i = 131$	$i = 134$	$i = 136$	
Odessa troilite									
200				116 $\pm$ 3	15.0 $\pm$ 0.3	80.7 $\pm$ 2.2	39.3 $\pm$ 1.3	33.7 $\pm$ 0.9	2.89
400	0.361 $\pm$ 0.081	0.343 $\pm$ 0.051	7.31 $\pm$ 0.26	134 $\pm$ 2	14.8 $\pm$ 0.4	77.9 $\pm$ 1.0	38.1 $\pm$ 0.6	32.9 $\pm$ 0.5	1.31
600	0.546 $\pm$ 0.111	0.286 $\pm$ 0.173	7.34 $\pm$ 0.56	158 $\pm$ 3	14.5 $\pm$ 0.3	77.3 $\pm$ 1.6	38.1 $\pm$ 0.7	32.1 $\pm$ 0.5	0.57
800	0.553 $\pm$ 0.214	0.364 $\pm$ 0.110	7.82 $\pm$ 0.57	239 $\pm$ 3	15.3 $\pm$ 0.7	82.5 $\pm$ 1.5	38.2 $\pm$ 0.7	38.8 $\pm$ 0.8	0.41
1,000	0.341 $\pm$ 0.050	0.332 $\pm$ 0.035	8.58 $\pm$ 0.12	325 $\pm$ 2	15.2 $\pm$ 0.2	94.1 $\pm$ 0.6	38.8 $\pm$ 0.4	33.2 $\pm$ 0.2	3.01
1,100	0.435 $\pm$ 0.042	0.385 $\pm$ 0.032	9.06 $\pm$ 0.15	424 $\pm$ 2	15.5 $\pm$ 0.2	80.2 $\pm$ 0.4	39.9 $\pm$ 0.2	33.9 $\pm$ 0.2	3.27
Cape York troilite									
200				101 $\pm$ 8					0.01
400				122 $\pm$ 10					0.01
600				107 $\pm$ 4	15.4 $\pm$ 0.9	82.5 $\pm$ 2.9	39.8 $\pm$ 1.2	34.4 $\pm$ 1.2	0.18
800			8.53 $\pm$ 0.86	127 $\pm$ 3	15.1 $\pm$ 0.4	88.3 $\pm$ 0.8	38.9 $\pm$ 0.8	34.0 $\pm$ 0.7	0.33
900			8.40 $\pm$ 1.25	171 $\pm$ 2	14.8 $\pm$ 0.6	100.3 $\pm$ 2.1	40.2 $\pm$ 1.1	34.2 $\pm$ 0.5	0.17
1,000			8.97 $\pm$ 1.89	144 $\pm$ 7	15.2 $\pm$ 1.2	91.4 $\pm$ 3.7	39.8 $\pm$ 1.1	33.7 $\pm$ 1.3	0.20
1,100	0.384 $\pm$ 0.037	0.385 $\pm$ 0.060	7.30 $\pm$ 0.38	161 $\pm$ 1	15.2 $\pm$ 0.4	93.9 $\pm$ 1.2	38.9 $\pm$ 0.7	33.2 $\pm$ 0.6	0.88
Air	0.357	0.335	7.14	98.3	15.17	78.8	38.8	33.0	
AVCC	0.459	0.410	8.20		16.1	81.7	38.2	32.1	

The sample of Cape York troilite was a single piece weighing 2.3236 g and that of Odessa troilite consisted of three pieces weighing 1.5429 g that had been taken from a single nodule.

To compare the isotopic composition of Xe in these two temperature fractions with that seen in chondrites¹⁸, the abundance of each stable isotope of mass, i , is defined by the ratio, R^i , where

$$R^i = (^i\text{Xe}/^{132}\text{Xe})_{\text{sample}} / (^i\text{Xe}/^{132}\text{Xe})_{\text{AVCC}} \quad (3)$$

Values of R^i are shown in Fig. 2 for each xenon isotope in the 1,000 °C fraction of Odessa troilite and in Fig. 3 for those in the 1,100 °C fraction of Cape York troilite.

In both gas fractions, the abundance pattern of the nonradiogenic isotopes matches that of terrestrial Xe. The high value of R^{131} indicates the release of neutron-capture-produced ^{131}Xe , and the high value of R^{128} in the 1,000 °C fraction of Odessa troilite (Fig. 2) probably indicates the release of ^{128}Xe from the $^{127}\text{I}(\text{n}, \gamma\beta^-)^{128}\text{Xe}$ reaction. The decay product of ^{129}I indicated by $^{129}\text{Xe}/^{132}\text{Xe} = 3.25$ in the 1,000 °C fraction of Odessa troilite also suggests the degassing of sites that contain iodine. A small contribution of spallation-produced Xe may be responsible for the slight enrichment of the light xenon isotopes shown in Fig. 3 for the Cape York troilite.

For xenon released from the troilite samples at most other extraction temperatures, Table 1 shows that values of the $^{136}\text{Xe}/^{132}\text{Xe}$ and $^{134}\text{Xe}/^{132}\text{Xe}$ ratios are generally higher than those of AVCC Xe and that values of the $^{130}\text{Xe}/^{132}\text{Xe}$ ratio are generally lower than that of AVCC Xe. The values of these three ratios are more like those in terrestrial than in AVCC Xe.

Values of the $^{129}\text{Xe}/^{133}\text{Xe}$, $^{128}\text{Xe}/^{132}\text{Xe}$ and $^{131}\text{Xe}/^{132}\text{Xe}$ ratios do not necessarily indicate the isotopic signature of the indigenous component because of *in situ* products from the decay of ^{129}I and from $(\text{n}, \gamma\beta^-)$ reactions on ^{128}Te , ^{127}I and ^{130}Te .

There are large uncertainties on the measurements of ^{124}Xe and ^{126}Xe because of their low natural abundance, but measured values of the $^{126}\text{Xe}/^{132}\text{Xe}$ ratio are generally lower than in AVCC Xe. Xenon in the 600, 800 and 1,100 °C fractions of Odessa troilite suggest the possible presence of an excess ^{124}Xe component like that recently reported in the 550–650 °C and 1,400–1,500 °C fractions of troilite-rimmed chondrules and blocky sulphides of the Allende chondrite⁴. Large uncertainties in the measurements of these light isotopes make it impossible to distinguish between spallation products, excess ^{124}Xe and/or a component of AVCC-type Xe. The absence of significant amounts of excess ^{131}Xe in the 600, 800 and 1,100 °C fractions of Odessa suggests that ^{124}Xe -enriched xenon is not anyway representative of the bulk Xe trapped within the troilite.

U-Xe, a hypothetical type of xenon recently proposed by Pepin and Phinney²², is characterized by values of the $^{124-131}\text{Xe}/^{132}\text{Xe}$ ratios that are higher than those of AVCC Xe and by values of $^{134-136}\text{Xe}/^{132}\text{Xe}$ ratios that are lower than those of AVCC Xe. Obviously U-Xe cannot explain the isotopic

composition of Xe released from our samples (see Table 1).

To check the melting temperature of troilite under vacuum, a 0.165-g sample of natural terrestrial troilite was heated in an evacuated quartz tube. The troilite was broken from a large single crystal (purchased from Wards Natural Science Establishment). The temperature was increased in 100 °C increments, and melting and partial dissociation was observed at 1,100 °C. The degassing of Odessa troilite at 1,000 °C may indicate crystal imperfections or chemical impurities because Odessa was not exposed to the prolonged annealing of Cape York or terrestrial troilite^{20,21}. However, the uncertainty in the extraction temperature, $\approx \pm 50$ °C, may account for the apparent difference in the melting temperatures of troilite from the Cape York and Odessa meteorites.

The data shown in Figs 1–3 demonstrate the presence of terrestrial Xe trapped in troilite of the Odessa and Cape York iron meteorites. There is growing evidence that the protoplanetary nebula was chemically and isotopically zoned^{5,6} and that the cores of the four inner planets formed in the central

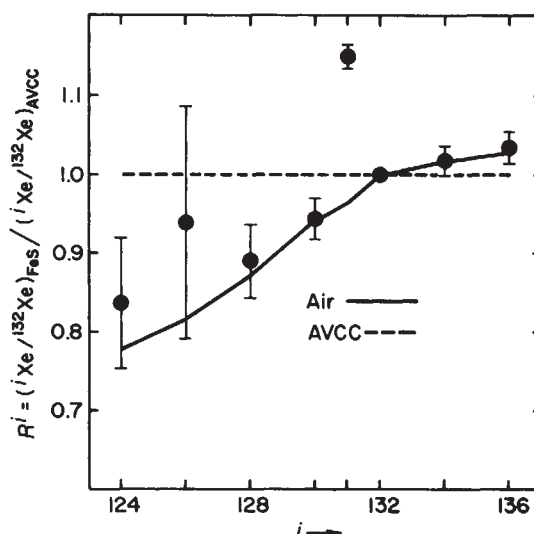


Fig. 3 A comparison of the isotopic composition of Xe in the 1,100 °C fraction of Cape York troilite (●) with that in air and in chondritic meteorites¹⁸, AVCC. The isotopic composition of Xe released from Cape York troilite with the tracer isotope of *in situ* generated ^{131}Xe is like that in air, except for a possible component of spallation-produced ^{124}Xe and ^{126}Xe . The isotopic composition of Xe trapped in meteoritic troilite differs from that seen in bulk chondrites¹⁸ across the entire spectrum of stable isotopes.

Fe-rich region^{5,6,23,24}. Since different mixes of nucleogenetic components have been identified in various phases of chondrites¹⁻⁵, the release of terrestrial Xe might also be expected in the melt fraction of troilite separates of chondrites. In fact, terrestrial Xe was reported recently⁴ in the gas released at 900 °C from troilite-rimmed chondrules (no. 20A1) and blocky iron sulphides of the primitive Allende carbonaceous chondrite. The observations were attributed to adsorbed atmospheric Xe, but this interpretation is difficult to reconcile with the release of distinctly non-terrestrial Xe at lower extraction temperatures⁴. Further, the amount of Xe released from the blocky sulphides and the troilite-rimmed chondrules at 900 °C exceeds that released in any other temperature fraction below 1,250 °C. Uncertainties in the extraction temperatures are not given for the Allende study⁴, but the pronounced release of Xe for the temperature fraction labelled 900 °C suggests that this corresponds to the degassing of Xe trapped in, rather than adsorbed on, Allende sulphides.

In air the ⁸⁶Kr/¹³²Xe ratio is 8.5, a factor of about 20 higher than that in carbonaceous chondrites. Lower values of Kr/Xe are observed in terrestrial rocks and minerals, and it is uncertain whether preferential retention of Xe by solids can account for its apparent depletion in air⁶. In the troilite samples from Odessa and Cape York the ⁸⁶Kr/¹³²Xe ratio is 1.6 and 2.0, respectively, a factor of about 4.7 higher than the ⁸⁶Kr/¹³²Xe ratio in carbonaceous chondrites. The higher Kr/Xe ratio in troilite may indicate differences in the value of this ratio in different parts of the protoplanetary nebula or differences in the degree of elemental fractionation by the processes that ultimately incorporated Kr and Xe into carbon and troilite.

Regardless of the mechanism that produced the high Kr/Xe ratio in air, the results shown in Figs 1–3 imply that the isotopic composition of terrestrial Xe represents the particular mix of nucleogenetic components that was dominant in that part of the protoplanetary nebula which contained high concentrations of sulphur and iron. Support for this conclusion can be found in earlier reports of a spike of terrestrial Xe released at ≈900–1,100 °C from iron sulphides of chondrites^{4,8} and iron meteorites⁹.

This research was supported by Weldon Springs Funds of the University of Missouri. We thank Professor Vagn F. Buchwald for providing the sample of Cape York used in this study, Mr Mark Stein for updating our mass spectrometer so that our analyses could be obtained rapidly and reliably by magnetic peak-jumping and computer data acquisition, and Mr Rajeswara Duppatla for writing the computer program used for the acquisition and reduction of output data.

Received 26 April; accepted 11 August 1982.

- Manuel, O. K., Hennecke, E. W. & Sabu, D. D. *Nature* **240**, 99–101 (1972).
- Frick, U. *Proc. 8th Lunar Sci. Conf.* 273–292 (1977).
- Srinivasan, B. & Anders, E. *Science* **201**, 51–56 (1978).
- Lewis, R. S., Hertogen, J., Alaerts, L. & Anders, E. *Geochim. cosmochim. Acta* **43**, 1743–1752 (1979).
- Sabu, D. D. & Manuel, O. K. *Meteoritics* **15**, 117–138 (1980).
- Manuel, O. K. & Sabu, D. D. *Geochim. J.* **15**, 245–267 (1981).
- Niemeyer, S. & Leich, D. A. *Proc. 7th Lunar Sci. Conf.* **2**, 587–597 (1976).
- Merrihue, C. J. *geophys. Res.* **71**, 263–313 (1966).
- Niemeyer, S. *Geochim. cosmochim. Acta* **43**, 843–860 (1979).
- Munk, M. N. *Earth planet. Sci. Lett.* **3**, 133–138 (1967).
- Hennecke, E. W. & Manuel, O. K. *Earth planet. Sci. Lett.* **36**, 29–43 (1977).
- Rowe, M. W. & Bogard, D. D. *J. geophys. Res.* **71**, 4183–4188 (1966).
- M. N. Munk, *Earth planet. Sci. Lett.* **3**, 457–465 (1967).
- Edgerley, D. A. & Rowe, M. W. *Geochim. J.* **13**, 103–112 (1979).
- Lightner, B. D. & Marti, K. *Proc. 5th Lunar Sci. Conf.* **2**, 2023–2031 (1974).
- Srinivasan, B., Alexander, E. C. Jr & Manuel, O. K. *J. inorg. Nucl. Chem.* **34**, 2381–2396 (1972).
- Nier, A. O. *Phys. Rev.* **79**, 450–454 (1950).
- Eugster, O., Eberhardt, P. & Geiss, J. *J. geophys. Res.* **71**, 3874–3896 (1969).
- Hwaung, C.-Y. G. thesis, Univ. Missouri-Rolla (1982).
- Wasson, J. T. *Geochim. cosmochim. Acta* **34**, 957–964 (1970).
- Scott, E. R. D., Wasson, J. T. & Buchwald, V. F. *Geochim. cosmochim. Acta* **37**, 1957–1983 (1973).
- Pepin, R. O. & Phinney, D. *Components of Xenon in the Solar System* (Preprint, Univ. Minnesota-Minneapolis, 1980).
- Turekian, K. K. & Clarke, R. S. Jr *Earth planet. Sci. Lett.* **6**, 346–348 (1969).
- Vinogradov, A. P. *Geokhimiya* **10**, 1427–1431 (1975).

Phase transition in KOH-doped hexagonal ice

Y. Tajima, T. Matsuo & H. Suga

Department of Chemistry and Chemical Thermodynamics Laboratory, Faculty of Science, Osaka University, Toyonaka, Osaka 560, Japan

A few crystals exist which have residual entropy^{1,2}. The most notable of these is hexagonal ice, *I_h*, the ordinary form of solid H₂O (ref. 3). The structural interpretation of the residual entropy proposed by Pauling and widely accepted ascribes it to positional disorder of the protons in the ice conditions⁴. As such a disordered arrangement of the constituent atom cannot be an equilibrium structure of the crystal at the lowest temperature, there has been great interest in the search for a possible ordering phenomenon in ice crystals. We report here a calorimetric experiment which shows that a first order phase transition takes place in KOH-doped ice crystals and that the transition removes most of the residual entropy of the ice crystal.

The presence of residual entropy implies a lack of thermal equilibrium in the crystal below a certain temperature. In fact, we have observed relaxational heat-capacity anomalies of ~100 K for ordinary ice⁵ and ~115 K for heavy ice⁶, and have studied the associated enthalpy relaxation by subjecting the ice specimen to rapid quenching and long annealing in an adiabatic calorimeter. The study revealed the relaxational behaviour to be a kind of glass transition in the stable crystal, thus eliminating the previous interpretation⁷ of the corresponding anomalies in dielectric properties as due to the onset of a ferroelectric phase transition. The enthalpy relaxation time became 10⁴ s at 105 K

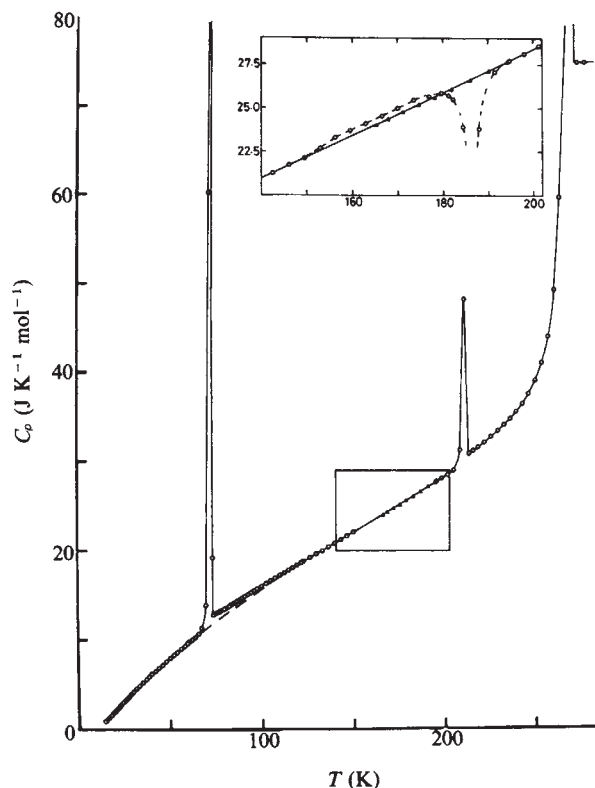


Fig. 1 The heat capacity of H₂O doped with 0.1 mol dm⁻³ KOH. The inset shows the effect of crystallization of the undercooled part of the sample.

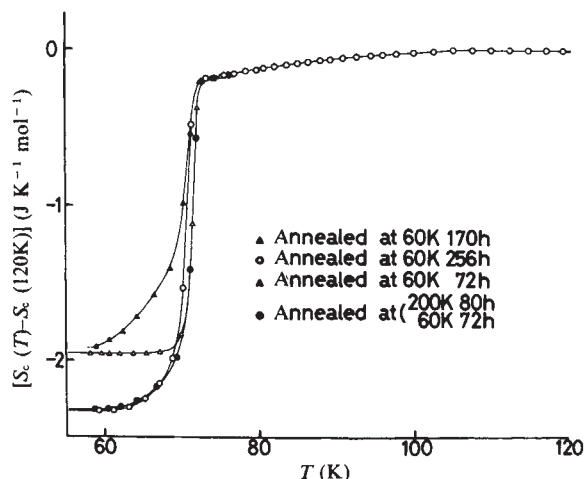


Fig. 2 The anomalous entropy of H_2O doped with 0.1 mol dm^{-3} KOH referred to 120 K. The key follows the order of the measurement.

and 10^6 s at about 87 K. We found that a completely ordered state at the hypothetical transition temperature, 60 K as calculated by Pitzer and Polissar⁸, would require an extremely long time.

Following the proposal by Onsager⁹, we have made heat-capacity measurements on HF-doped ice crystals¹⁰. Doping by this impurity is thought to introduce various kinds of lattice defects, including the Bjerrum fault¹¹, and to accelerate the rate of rearrangement of proton configurations by catalysing proton migration. However, our heat-capacity measurements actually revealed a positive effect of HF-doping, in that it increased the heat-capacity anomaly and thus brought about a more ordered state. The relaxation time was reduced by a factor of 30 at the relevant temperature for the most effective concentration of HF dissolved in the ice specimen, but not enough for a phase transition to take place.

When we doped an ice sample with a small amount of KOH, however, a drastic change occurred in its thermal behaviour. Aqueous solutions of KOH of the desired concentration were prepared by stepwise dilution. The doped water specimen was sealed in a Mylar (polyethylene terephthalate film) bag and put in a calorimeter cell together with 1 atm pressure of He gas as a heat exchanger. The double-jacketed adiabatic calorimeter operating in an intermittent heating mode used in the present study has been described previously¹². The calorimeter cell was made of gold-plated copper and had an internal volume of 8 cm^3 . The temperature of the cell was measured by a platinum resistance thermometer (Minco s1059, Minco Products) along with a precision a.c. double bridge (Automatic System Laboratory Inc., model H8). Resolution of the temperature measurement was better than 0.1 mK in the relevant temperature region. The precision of the heat-capacity data was generally better than 0.1%.

The calorimeter containing 0.1 mol dm^{-3} KOH solution was cooled in the cryostat. The time required for the completion of the crystallization of the 5.5406 g sample was about 15 min as indicated by the duration of the temperature arrest at 273 K. The crystallized sample was cooled to 80 K at 2 K min^{-1} and then to 60–65 K. It was then kept in an adiabatic condition and its temperature recorded. The temperature of the cell increased spontaneously for about 1 week. When the heat evolution ceased, the calorimetric system was cooled to 13 K and the heat-capacity measurement was started. The heat-capacity data are given in Fig. 1 and show three anomalies, at 72, 208 and 273 K. The second anomaly is due to the eutectic melting of the ice and hydrated KOH and the third to the melting of ice into the KOH solution. The anomaly at 72 K is due to the phase transition and is essentially isothermal. The transition

occurs neither in pure ice nor in H_2O doped with HF. The excess heat capacity of the KOH-doped ice over that of pure ice extends over a wide temperature range, up to 40 K above the transition temperature. It is this high-temperature tail of excess heat capacity that we have observed as the relaxational heat capacity in the pure ice crystal. We chose 120 K as the reference temperature for calculating the difference in the entropies between the pure and KOH-doped ices from the heat-capacity data. As Fig. 2 shows, the decrease of the entropy in the phase transition is $2.32 \text{ J K}^{-1} \text{ mol}^{-1}$. Thus, the phase transition removes a substantial fraction ($\approx 70\%$) of the configurational entropy that the ice crystal retains at higher temperatures. In a similar experiment on D_2O the phase transition occurred at 76 K.

As shown in Fig. 2, the magnitude of the entropy removed by the transition depended in a complex manner on the temperature and time of annealing as well as on the previous history of the sample, for example, whether or not it had undergone the phase transition before, or the temperature to which it had been heated after the previous phase transition. However, we simply record here the occurrence of the hysteretic behaviour without going into its analysis, because we are primarily concerned with the determination of the total transition entropy.

We believe that the anomaly is an intrinsic equilibrium property of ice I_h which has previously escaped observation because of its kinetics and which has now been revealed by the catalytic action of KOH. The reasons for this are: (1) the temperature of the heat-capacity peak is independent of the concentration of KOH. In experiments using 10 and 100 times lower concentrations of KOH solution, the transition occurred at the same temperature, as Fig. 3 shows for the $0.001 \text{ mol dm}^{-3}$ specimen. (2) The small amount of KOH (≤ 0.002 by the amount-of-substance fraction) is not expected to change the bulk equilibrium property of the ice to this extent. However, this statement will be tested further by experiments using different dopants such as LiOH, RbOH and $\text{Ba}(\text{OH})_2$.

Use of a polycrystalline sample is justified by the large entropy change of the transition observed, as surface contribu-

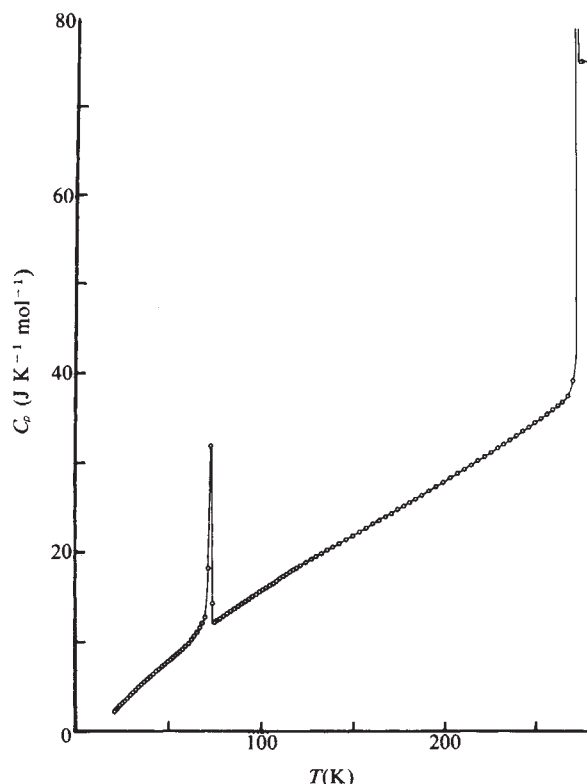


Fig. 3 The heat capacity of H_2O doped with $0.001 \text{ mol dm}^{-3}$ KOH.

tions to the molar entropy for a reasonable crystallite size are much smaller than the observed transition entropy. Both poly crystals and single crystals have been used in calorimetry without any specific surface effects^{3,5,6,13,14}. The crystallite size may influence the rate of the phase transition through the size dependence of the effective dopant concentration in the bulk of the specimen and through surface nucleation of the new phase.

The solubility of HF in ice and characterization of the defects introduced by it have been reported previously^{15,16}. However, very little is known about KOH impurities in ice. The present study shows that the ions are incorporated in the lattice. It would be useful to investigate the properties of the dissolved ions by microscopic methods.

Kawada studied the dielectric properties of doped ices and found a phase transition at 70 K on a KOH-doped specimen¹⁷. The temperature of the dielectric phase transition agrees with the present observation. Theoretical estimates of the transition temperature, depending to different extents on the experimental data, vary between 60 and 73 (refs 8, 18, 19). The agreement between the experimental and theoretical transition temperatures shows that the electrostatic models underlying these calculations are essentially correct. Structural investigation of the low temperature phase in collaboration with a neutron group is being planned.

This is contribution no. 44 from the Chemical Thermodynamics Laboratory.

Received 18 May; accepted 3 August 1982.

1. Wilkes, J. *The Third Law of Thermodynamics* (Oxford University Press, 1961).
2. Parsonage, N. G. & Staveley, L. A. K. *Disorder in Crystals* (Clarendon, Oxford, 1978).
3. Giauque, W. F. & Stout, J. W. *J. Am. chem. Soc.* **58**, 1144–1150 (1936).
4. Pauling, L. *J. Am. chem. Soc.* **57**, 2680–2684 (1935).
5. Haida, O., Matsuo, T., Suga, H. & Seki, S. *J. chem. Thermodyn.* **6**, 815–825 (1974).
6. Haida, O., Suga, H. & Seki, S. *J. Glaciol.* **22**, 155–164 (1979).
7. Dengel, O., Eckener, V., Plitz, H. & Riehl, N. *Phys. Lett.* **9**, 291–293 (1964).
8. Polissar, J. & Pitzer, K. S. *J. phys. Chem.* **60**, 1140–1142 (1956).
9. Onsager, L. in *Ferroelectricity* (ed. Weller, E.) 16–19 (Elsevier, Amsterdam, 1967).
10. Ueda, M., Matsuo, T. & Suga, H. *J. Phys. chem. Solids* (in the press).
11. Bjerrum, N. *Science* **115**, 385–390 (1952).
12. Tatsumi, M., Matsuo, T., Suga, H. & Seki, S. *Bull. chem. Soc. Japan* **48**, 3060–3066 (1975).
13. Van den Beukel, A. *Phys. Stat. Solids* **28**, 565–568 (1968).
14. Pick, M. A., Wenzel, H. & Engelhardt, H. *Z. Naturforsch.* **26A**, 810–814 (1971).
15. Haltenorth, H. & Klinger, J. *J. Solid St. Commun.* **21**, 523–535 (1977).
16. Bilgram, J. H., Roos, J. & Gränicher, H. *Z. Phys.* **B23**, 1–9 (1976).
17. Kawada, S. *J. phys. Soc. Japan* **32**, 1442 (1972).
18. Hentschel, H. G. *Molec. Phys.* **38**, 401–411 (1979).
19. Minagawa, I. *J. phys. Soc. Japan* **50**, 3669–3676 (1981).

Influence of cross-diffusion on 'finger' double-diffusive convection

Trevor J. McDougall & J. Stewart Turner

Research School of Earth Sciences, Australian National University, PO Box 4, Canberra 2600, Australia

Structures similar to the oceanic 'salt fingers' that form when warmer more saline water overlies cooler fresher water¹ have recently been observed², in concentrated solutions of polymers in conditions where fingers would not form according to the usual criteria. We here extend previous theories to derive the conditions under which double-diffusive convection can occur in a solution of a pair of solutes with a large coupled diffusion (or cross-diffusion) effect, that is, where a flux of one property is driven by a spatial gradient of another.

Salt fingers have an important role in the mixing of properties in several regions of the ocean³. The fingers form a pattern in which each upward-moving finger is surrounded by downward-moving fingers and vice versa. The downgoing fingers continually lose heat (by horizontal conduction) to the neighbouring upgoing fingers, so making the downgoing fingers more dense

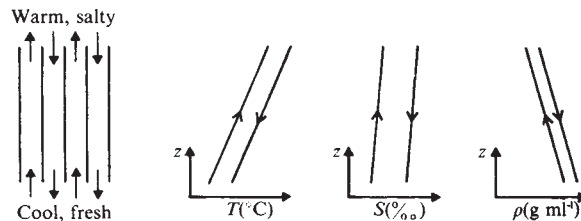


Fig. 1 Up- and down-going salt fingers and their temperatures, salinities and densities. The arrows on the individual profiles distinguish the up from the downgoing fingers. Note that in contrast to Fig. 3, the temperature is plotted here rather than its effect on density.

and the upgoing fingers less dense (Fig. 1). The small width of the fingers (shown by both theory and direct observations to be a few centimetres) allows a significant heat flux to occur by molecular heat conduction, leading to the temperature gradients in the fingers shown in Fig. 1. Because the molecular diffusivity of salt is much less than that of heat, the fingers do not transfer much salt between themselves, so the vertical salinity gradient is small (Fig. 1). Figure 2 is a cross-section through a field of fingers which are growing downwards into a stable temperature gradient set up in a laboratory tank.

Preston *et al.*² have recently produced striking photographs of fingers in studies of the diffusion of macromolecules. In their paper three polymers were used, but in their subsequent work⁴ they have observed fingers at a horizontal interface between an upper dextran solution and a lower, more concentrated dextran solution which also has a small amount of sorbitol added to it. This configuration is not obviously conducive to fingering because both solutes (dextran and sorbitol) are more concentrated in the lower layer. We show here that with a sufficiently large cross-diffusion effect, fingers can form in such a situation.

It is not uncommon, even today, to read in the diffusion literature that convective motion cannot occur at, say, a diffusing horizontal interface between two solutions unless the vertical density gradient ρ_z is positive, that is, unless the fluid is statically unstable. Both types of double-diffusive convection occur, however, in a statically stable density gradient, where $\rho_z < 0$. Wendt⁵ has solved the problem of the diffusion of two properties from an initially sharp interface with two cross-diffusion terms and has shown that such one-dimensional diffusion can lead to a local reversal of the density gradient. This result has been widely used to determine whether convection can be expected in particular situations, often associated with measurements of diffusion coefficients (in, for example, a Gouy diffusimeter) and the testing of the Onsager relations.

Double-diffusive convection developed in the 1960s mainly due to its applications in oceanography but recently it has been realized that it can be important in many other contexts¹. Two recent papers^{6,7} have investigated double-diffusive instabilities



Fig. 2 A field of salt fingers formed by setting up a stable temperature gradient and pouring a little salt solution on top. The downward-moving fingers are made visible by adding fluorescein to the salt solution and lighting through a slit from below.

caused by the Soret effect in the special geometry of a Rayleigh-Bénard convection experiment where a solution is differentially heated from above and cooled below, showing that convective instabilities can and do occur when the density gradient is statically stable. Hurle and Jakeman⁸ have studied the case where the heating and cooling are reversed with the heating occurring at the bottom and the cooling above.

We have incorporated the two cross-diffusion flux terms into the linear stability analysis of double-diffusive convection and we focus our attention on the conditions under which fingers form when *both* components are stably distributed. We consider a solution of two solutes T and S confined between two horizontal porous plates which maintain a fixed contrast in fluid properties between the plates. We parameterize the diffusion and the cross-diffusion fluxes by

$$\text{--Flux of } T = D_{11}\nabla T + D_{12}\nabla S \quad (1)$$

$$\text{--Flux of } S = D_{22}\nabla S + D_{21}\nabla T \quad (2)$$

The density of the fluid ρ is given by the linear relation

$$\rho = \rho_0(1 + \alpha T + \beta S) \quad (3)$$

where α and β are constants, ρ_0 is a reference density and T and S are the concentrations of the two solutes, defined so that $D_{11} > D_{22}$. That is, T has a larger molecular diffusivity than S and we define the ratio D_{22}/D_{11} to be equal to $\tau (< 1)$. On following through the linear stability analysis for small perturbations from linear property gradients and no initial motion, it can be shown⁹ that the finger instability will occur if

$$\left(\frac{\beta D_{21}}{\alpha D_{22}} - 1\right) + \frac{1}{\tau} \frac{\beta S_z}{\alpha T_z} \left(\frac{\alpha D_{12}}{\beta D_{11}} - 1\right) > \frac{27}{4} \frac{\pi^4}{R} \left(1 - \frac{D_{12}D_{21}}{D_{11}D_{22}}\right) \quad (4)$$

This inequality is appropriate to the case where T is distributed in a statically stable fashion (that is if $\alpha T_z < 0$, where z is positive upwards and the subscript denotes differentiation). Here R is the Rayleigh number defined in the usual way and normally it is large in comparison to $27\pi^4/4$ and so the right side of inequality 4 can be taken as zero.

Inequality 4 is the criterion which must be satisfied if fingers are to form when the *in situ* vertical gradients of T and S are T_z and S_z (assumed uniform in the horizontal plane). The normal condition for fingers to form without cross-diffusion is readily recovered from inequality (4). In this case, $D_{12} = D_{21} = 0$, $\beta S_z > 0$, $\alpha T_z < 0$, giving the well known condition $|\beta S_z/\alpha T_z| > \tau$. Hydrostatic instability ($\rho_z > 0$) occurs when $|\beta S_z/\alpha T_z| > 1$ and so this requires a considerably larger gradient of S than is required for the formation of fingers. We postpone further discussion of inequality (4) until after we examine a finite amplitude finger model.

Huppert and Manins¹⁰ have presented a linear model of finite-amplitude steady, very long salt fingers. The nonlinear advective terms in the Navier-Stokes equations are equal to zero because the motion is assumed to be entirely vertical. The vertical velocity W and the deviations T' and S' of the T and S properties from their horizontal averages are assumed to take the form

$$(W, T', S') = (W_0, T_0, S_0) \cos mx \cos ny \quad (5)$$

where x and y are the horizontal coordinates. By substituting equation(s) (5) into the Navier-Stokes equations (significantly simplified by the assumption of purely vertical velocities) we obtain

$$\begin{aligned} & -\alpha \bar{T}_z \left(\frac{\beta}{\alpha} D_{21} - D_{22} \right) - \beta \bar{S}_z \left(\frac{\alpha}{\beta} D_{12} - D_{11} \right) \\ & = \frac{\nu k^4}{g} (D_{11}D_{22} - D_{12}D_{21}) \end{aligned} \quad (6)$$

where $k^2 = m^2 + n^2$. For fingers to form, k^4 must be positive,

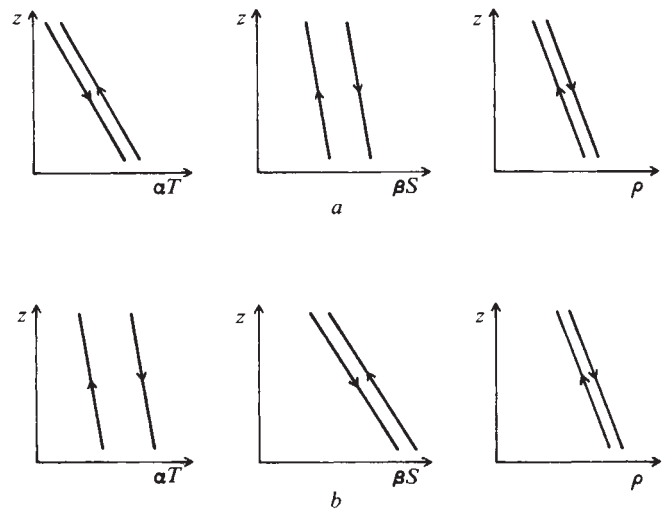


Fig. 3 Sketch of the vertical profiles of αT , βS and ρ in the centres of upgoing and downgoing (indicated by the arrows) fingers for $\bar{T}_z < 0$, $\bar{S}_z < 0$; that is with both T and S stably stratified. Panel *a* has $D_{12} = 0$ and D_{21} satisfies inequality (7) (see text). Panel *b* has $D_{21} = 0$ and D_{12} satisfies inequality (7).

and taking $D_{11}D_{22} > D_{12}D_{21}$ we see that the left hand side of equation (6) must be positive. $\bar{T}_z$ and $\bar{S}_z$ are the vertical gradients of the horizontal average of T and S through the fingers. We concentrate on the case where T is stably stratified ($\alpha T_z < 0$) and from equation (6) we see that finite amplitude fingers will persist if

$$\left(\frac{\beta D_{21}}{\alpha D_{22}} - 1\right) + \frac{1}{\tau} \frac{\beta \bar{S}_z}{\alpha \bar{T}_z} \left(\frac{\alpha D_{12}}{\beta D_{11}} - 1\right) > 0 \quad (7)$$

This inequality looks very similar to inequality (4) but the two relations are derived by very different theoretical methods. Inequality (4) is the result of a linear stability analysis for the onset of fingers in uniform vertical property gradients T_z and S_z , while inequality 7 is the condition for the maintenance of finite amplitude fingers. Apart from the different interpretation of the vertical property gradients in inequalities (4) and (7) the two criteria are the same.

Consider now the situation where both T and S are stably stratified, that is $\alpha T_z < 0$ and $\beta S_z < 0$. We see from inequality 7 that the occurrence of fingers is assisted by large positive values of both D_{12} and D_{21} and that at least one of $\beta D_{21}/\alpha D_{22}$ and $\alpha D_{12}/\beta D_{11}$ must exceed unity.

The driving energy for normal fingering motion comes from the unstably stratified component S and this gravitational potential energy is released by the rapid diffusion of T between the fingers. In the fingering motion we have just described we have shown that even if both components, T and S are distributed in a hydrostatically stable fashion, finger convection can still occur if the cross-diffusion coefficients D_{12} and D_{21} are large enough to satisfy inequality (7). The total gravitational potential energy of the fluid still decreases due to the fingering, but in this case it is the cross-diffusion between the fingers which allows the release of gravitational potential energy even though no separate solute is unstably stratified.

The T , S and ρ profiles in upgoing and downgoing fingers are shown in Fig. 3 for the two special cases when one of the cross-diffusion coefficients is zero. In Fig. 3a where $D_{12} = 0$, the T concentration of the downgoing fingers increases due to the diffusion of T from the upgoing into the downgoing fingers. This flux between the fingers occurs in the normal fashion because the T flux is unaffected by coupled diffusion. The flux of S between the fingers is however mainly due to the cross-diffusion flux of S caused by the horizontal spatial gradients of T . Similarly in Fig. 3b where $D_{21} = 0$, the flux of S between the

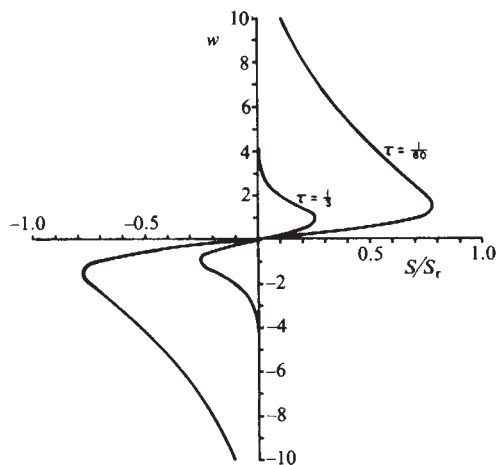


Fig. 4 Vertical profiles of the S concentration at a diffusing interface with $D_{12}=0$ and with the initial contrast in S concentration between the two layers equal to zero. The vertical coordinate w is equal to $z/2\sqrt{D_{22}t}$ and the reference S value S_r is $(\Delta T/2)(D_{21}/D_{22})\tau/(1-\tau)$.

fingers is not affected by coupled diffusion and the flux of T between the fingers is mainly due to the cross-diffusive flux of T caused by the positive value of D_{12} (satisfying inequality (7)) and the S gradient between the fingers. Note that in both a and b of Fig. 3 the horizontally averaged density gradient is statically stable ($\bar{\rho}_z < 0$) and that the downgoing fingers are more dense than the surrounding upgoing fingers.

The theoretical arguments outlined above have all related to regions in which there are constant vertical gradients of the two diffusing properties. A more common experimental situation, however, is one in which a layer of fluid is placed above another layer of different composition, with initially a sharp interface between them. When two solutions are placed one above the other, the diffusion of the two species T and S across the bounding interface may set up gradients which are favourable for the formation of fingers. Figure 4 shows the profiles of S as a function of $w = z/2\sqrt{D_{22}t}$ when the difference in S concentration between the layers is zero and one of the cross-diffusion coefficients, D_{12} , is also equal to zero. The S gradient is proportional to the cross diffusion coefficient D_{21} and to the difference in the T concentrations between the layers, ΔT . The S gradient is conducive to fingering in this interfacial region where the T gradient maintains the hydrostatic stability. This subject of diffusion at an initially sharp interface is discussed in more detail in ref. 9, including cases where both properties are initially set up in the hydrostatically 'stable' sense.

To test our belief that fingers can form due to cross-diffusion without needing to invoke any other special properties of polymers, we initiated experiments using two layers of sodium chloride solution at different temperatures, the cross diffusion term being now the Soret coefficient. According to published values of the coefficients, the criterion for instability to fingering motions (inequality (4)) could just be satisfied at low temperatures (where D_{21} is positive), with the upper layer a few degrees warmer than the lower. This system was, however, too sensitive to extraneous effects such as heating at the side walls and evaporation at the surface, and we concluded that a definitive experimental test of the theoretical ideas presented here will have to be carried out in a solute-solute system.

Cussler¹¹ lists references to measurements of all four ternary diffusion coefficients in various systems, showing that one of the few substances to exhibit large cross-diffusion fluxes is another polymer, polystyrene, of molecular weight of 190,000. Cussler and Lightfoot¹² used the polystyrene-toluene-cyclohexane combination and found values of $D_{12}/D_{11} < -1$. They also give a simple explanation of the large cross-diffusion effect in terms of the chemical potential as a function of temperature

for the two binary mixtures, polystyrene-toluene and polystyrene-cyclohexane.

In order to understand the mechanism of finger formation in situations such as those discussed by Preston *et al.*^{2,4} we suggest that the next logical step is to endeavour to measure the cross-diffusion fluxes for some of these long chain polymers in the presence of each other.

Received 7 May; accepted 4 August 1982.

1. Huppert, H. E. & Turner, J. S. *J. Fluid Mech.* **106**, 299–329 (1981).
2. Preston, B. N., Laurent, T. C., Comper, W. D. & Checkley, G. *J. Nature* **287**, 499–503 (1980).
3. Williams, A. J. *J. geophys. Res.* **86**, 1917–1928 (1981).
4. Comper, W. D. (in preparation).
5. Wendt, R. P. *J. phys. Chem.* **66**, 1740–1742 (1962).
6. Schechter, R. S., Prigogine, I. & Hamm, J. R. *Phys. Fluids* **15**, 379–386 (1972).
7. Velarde, M. G. & Schechter, R. S. *Phys. Fluids* **15**, 1707–1714 (1972).
8. Hurlle, D. T. J. & Jakeman, E. *J. Fluid Mech.* **47**, 667–687 (1971).
9. McDougall, T. J. *J. Fluid Mech.* (in the press).
10. Huppert, H. E. & Manins, P. C. *Deep-Sea Res.* **20**, 315–323 (1973).
11. Cussler, E. L. *Multicomponent Diffusion* (Elsevier, New York, 1976).
12. Cussler, E. L. & Lightfoot, E. N. *J. phys. Chem.* **69**, 1135–1144 (1965).

Cyanophyte calcification and changes in ocean chemistry

Robert Riding

Department of Geology, University College, Cardiff CF1 1XL, UK

Cyanophytes range from at least 2,200 Myr (ref. 1) to the Recent, but they only produced common marine shelly fossils during the Palaeozoic and Mesozoic (570–80 Myr) (Fig. 1). This contrasts with the pattern of metazoan evolution in which rapid diversification near the Precambrian–Cambrian boundary was closely accompanied by skeletonization² which has been retained in marine environments to the Recent. Attempts to explain this unusual geological distribution of marine calcareous cyanophytes cannot be made solely by reference to biological processes because these algae are mainly dependent on environmental conditions for their calcification³. Thus, the presence or absence of calcified cyanophytes may be a general indication of long-term changes in seawater chemistry. This likelihood has been recognized previously⁴ but has not been explored in any detail. Here I outline some possible explanations and suggest that cyanophyte calcification was facilitated by enhancement of marine CaCO_3 precipitation rates in the late Precambrian because of decrease in the $\text{Mg}^{2+}/\text{Ca}^{2+}$ ratio, linked to falling P_{CO_2} levels and extensive dolomite formation. Scarcity of calcareous cyanophytes in Cenozoic marine environments again implicates the $\text{Mg}^{2+}/\text{Ca}^{2+}$ ratio, inferred from oöid mineralogy to have increased in the late Mesozoic⁵, as a factor influencing cyanophyte calcification in the sea.

Calcareous cyanophytes, represented by microscopic fossils such as *Angulocellularia*, *Girvanella* and *Ortonella*, are consistently present in normal marine shelf environments from the Nemakit–Daldyn Horizon of latest Precambrian age in Siberia⁶ until the Jurassic and possibly also Cretaceous⁷. But they have not been reported from Cenozoic marine deposits and the very few records of them in Recent seas indicate that they are extremely rare, are not calcified to the same degree as Palaeozoic and Mesozoic examples, and tend to be limited to marginal marine situations^{8,9}. Similarly, calcareous cyanophytes appear to be absent from most Precambrian rocks despite intensive study of stromatolite microfabrics. The best evidence for Precambrian calcareous cyanophytes is weakly calcified traces of filaments from Upper Riphean and younger rocks in the Soviet Union⁶.

In attempting to account for the paucity of calcareous cyanophytes in the sea during the Precambrian and Cenozoic,

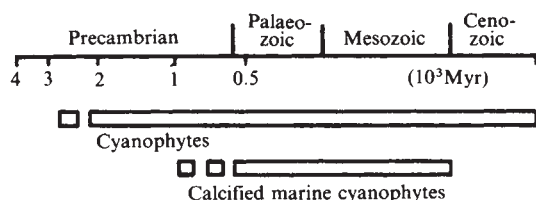


Fig. 1 Precambrian to Recent range of cyanophytes¹ compared with the Palaeozoic–Mesozoic occurrence of marine calcified cyanophytes.

both biological and environmental factors needed to be considered. The possibility that cyanophytes evolved calcification in the latest Precambrian cannot be ruled out, but silicified cyanophytes from the Proterozoic are similar to Recent forms¹⁰ and do not show morphological differences that could account for an absence of calcification. Furthermore, there is no good evidence that cyanophytes capable of calcification either became extinct or were excluded from marine environments in the late Mesozoic. The suggestion that cyanophytes in the sea were outcompeted by red and green algae⁴ is not supported by direct evidence^{11,12}, and it is known that Recent genera which calcify in freshwater can live uncalcified in the sea⁴. Consequently, an environmental control to account for marine cyanophyte calcification is not unreasonable, despite the fact that suggestions of changes in seawater composition near the Precambrian–Cambrian boundary to account for metazoan skeletonization have received little support recently². Mg^{2+}/Ca^{2+} ratio changes near the Precambrian–Cambrian boundary have been suggested in this context¹³, but metazoan skeletonization probably does not share the environmental dependence of cyanophyte calcification. However, it is remarkable that calcification in cyanophytes and metazoans began almost synchronously near 570 Myr.

Thermodynamically, an increase in temperature, partial pressure of CO_2 in buffered seawater, pH, or $a_{Ca^{2+}}$ will enhance the magnitude of $CaCO_3$ precipitation. A decrease in salinity resulting in increases of $\gamma_{Ca^{2+}}$ and $\gamma_{CO_3^{2-}}$ will have the same effect¹⁴. However, these parameters do not necessarily reflect the rate of precipitation of calcareous sediments. In marginal marine embayments it is possible to modify substantially temperature, pH and salinity, but these local changes in restricted environments cannot be invoked as long-term effects on general seawater chemistry. The occurrence of calcareous cyanophytes in reef environments from Cambrian to Jurassic–Cretaceous time⁷ requires modification of normal shelf seawater to be sustained for hundreds of millions of years. The rate of nucleation and growth of calcite, aragonite and dolomite, has been shown to depend on Mg^{2+} ion concentrations and a plausible explanation for changes in cyanophyte calcification can be invoked on the basis of kinetic considerations involving this ion.

The mutual interdependence of P_{CO_2} and the ratio of Mg^{2+}/Ca^{2+} in carbonated buffered seawater is well documented¹⁵. For example, an increase in P_{CO_2} will lead to $CaCO_3$ deposition and thus an increase of the Mg^{2+}/Ca^{2+} ratio. Conversely, any decrease in P_{CO_2} during the Precambrian¹⁶ would thus decrease the Mg^{2+}/Ca^{2+} ratio of seawater. As it has been demonstrated^{17,18} that relative decreases of Mg^{2+} concentrations greatly enhance the rate of nucleation of $CaCO_3$, it might be expected that cyanophyte calcification would be stimulated as a result. Extensive formation of late Precambrian dolomites¹⁹, suggested to be primary²⁰, would also have depressed the Mg^{2+} content of seawater. Reversal of this condition—increase of the Mg^{2+}/Ca^{2+} ratio—in the late Mesozoic has been attributed to widespread deposition of $CaCO_3$ by pelagic microorganisms as calcareous oozes, now preserved as chalks⁵. Hence, changes in the Mg^{2+}/Ca^{2+} ratio can be inferred which correlate with both the appearance of calcareous marine cyanophytes near the Precambrian–Cambrian boundary and their disappearance in the late Mesozoic.

Calcification may have conferred biological advantages on cyanophytes and certainly resulted in important changes in their sedimentological roles, whatever its cause may have been. An environmental control is plausible and implies that cyanophyte calcification could be an index of marine $CaCO_3$ precipitation rates which reflects times when inorganic carbonate production was increased. If this is so then it should also correlate with enhanced ooid formation and subsea cementation in the geological record. The presence of weakly calcified ?cyanophytes in the upper Proterozoic⁶ suggests a period of transition before the abrupt appearance of heavily calcified forms in the Nemakit–Daldyn. More information is required from the late Mesozoic to see whether a similar phase preceded the disappearance of calcareous marine cyanophytes at that time.

This research was supported by the NERC and also by the Humboldt Foundation through a Fellowship at the Technical University, Munich. I thank Peter Williams for advice on carbonate chemistry, and Robin Bathurst, Martin Brasier, Dianne Edwards, and Maurice Tucker for critically reading the manuscript.

Received 28 June; accepted 16 August 1982.

1. Cloud, P. *Paleobiology* **2**, 351–387 (1976).
2. Stanley, S. M. *Am. J. Sci.* **276**, 56–76 (1976).
3. Golubic, S. in *The Biology of Blue-Green Algae* (eds Carr, N. G. & Whitton, B. A.) 434–473 (Blackwell, Oxford, 1973).
4. Monty, C. L. V. *Ann. Soc. géol. Belg.* **96**, 585–624 (1973).
5. Sandberg, P. A. *Sedimentology* **22**, 497–538 (1975).
6. Riding, R. & Voronova, L. G. *Naturwissenschaften* (in the press).
7. Wray, J. L. *Developments in Palaeontology and Stratigraphy* Vol. 4 (Elsevier, Amsterdam, 1977).
8. Winland, H. D. & Matthews, R. K. *J. sedim. Petrol.* **44**, 921–927 (1974).
9. Golubic, S. & Campbell, S. E. in *Phanerozoic Stromatolites* (ed. Monty, C.) 209–229 (Springer, Berlin, 1981).
10. Schopf, J. W. *Scient. Am.* **239**, 85–102 (1978).
11. Gebelein, C. D. in *Developments in Sedimentology* Vol. 20 (ed. Walter, M. R.) 499–515 (Elsevier, Amsterdam, 1976).
12. Stanley, S. M. *Paleobiology* **2**, 209–219 (1976).
13. Durov, S. A. *Trudy novocherk. politekh. Inst.* **98**, (1960).
14. Garrels, R. M. & Christ, C. L. *Solutions, Minerals and Equilibria* (Harper & Row, New York, 1965).
15. Holland, H. D. *Proc. natn. Acad. Sci. U.S.A.* **53**, 1173–1182 (1965).
16. Berner, L. V. & Marshall, L. C. *Discuss. Faraday Soc.* **37**, 122–141 (1964).
17. Pytkowicz, R. M. *J. Geol.* **73**, 196–199 (1965).
18. Berner, R. A. *Soc. Econ. Paleont. Miner. Spec. Publ.* **20**, 37–43 (1974).
19. Ronov, A. B. *Geokhimiya* **8**, 715–743 (1964).
20. Tucker, M. E. *Geology* **10**, 1, 7–12 (1982).

Evidence for a central Eurasian source area of Arctic haze in Alaska

G. E. Shaw

Geophysical Institute, University of Alaska, Fairbanks, Alaska 99701, USA

During winter when the polar oceans are frozen, air masses entering Alaska from the Arctic are charged with suspended submicrometre particles whose chemical signatures show evidence of being derived from man-made sources of pollution^{1–3}. Occasionally, the aerosol loading is large enough to reduce visibility and thus the phenomenon has come to be referred to as 'Arctic haze'. We report here three strong episodes of Arctic haze in Alaska which were examined during February–April 1982 and which were found to be possibly associated with air emissions in central Eurasia.

Figure 1 illustrates the episodic nature of Arctic haze in interior Alaska during late winter 1982. The cross-hatched regions are synonymous with incidences of visible haze, which occurred in association with low air temperatures, increased particle concentration for particles near 0.13- μ m diameter (determined with a laser spectrometer) and, as the maps at the

bottom of Fig. 1 show, air trajectories (850 mbar level) coming from Arctic regions west of Alaska.

Particles sampled with impactors and examined with X-ray spectrometry during the time sequence represented in Fig. 1 were found to have very different X-ray spectra, depending on the direction which air masses had taken entering central Alaska. When the air flow had entered from the Gulf of Alaska or from the region of northwestern Canada, the visibility often exceeded 200 km; such times are referred to as 'clean times' and X-ray spectra of the minority (1%) non-sulphate particles taken then were primarily mixtures of Al and Si (crustal) or

ing the incoming air flow was anticyclonic activity over eastern Siberia and a trough over north-central Eurasia. The simultaneous occurrence of such flow conditions (which occur most frequently in the mid-winter months) and the enhanced metal X-ray signatures in aerosol suggest the presence of significant aerosol sources in central Eurasia, particularly around the area of the Taymyr Peninsula (there are no other apparent sources along the transport pathway). A Siberian source region of aerosol has also been deduced for the Canadian Arctic⁴.

Locating the exact geographical region or type of emission in central Eurasia from the back trajectories or from ratios of

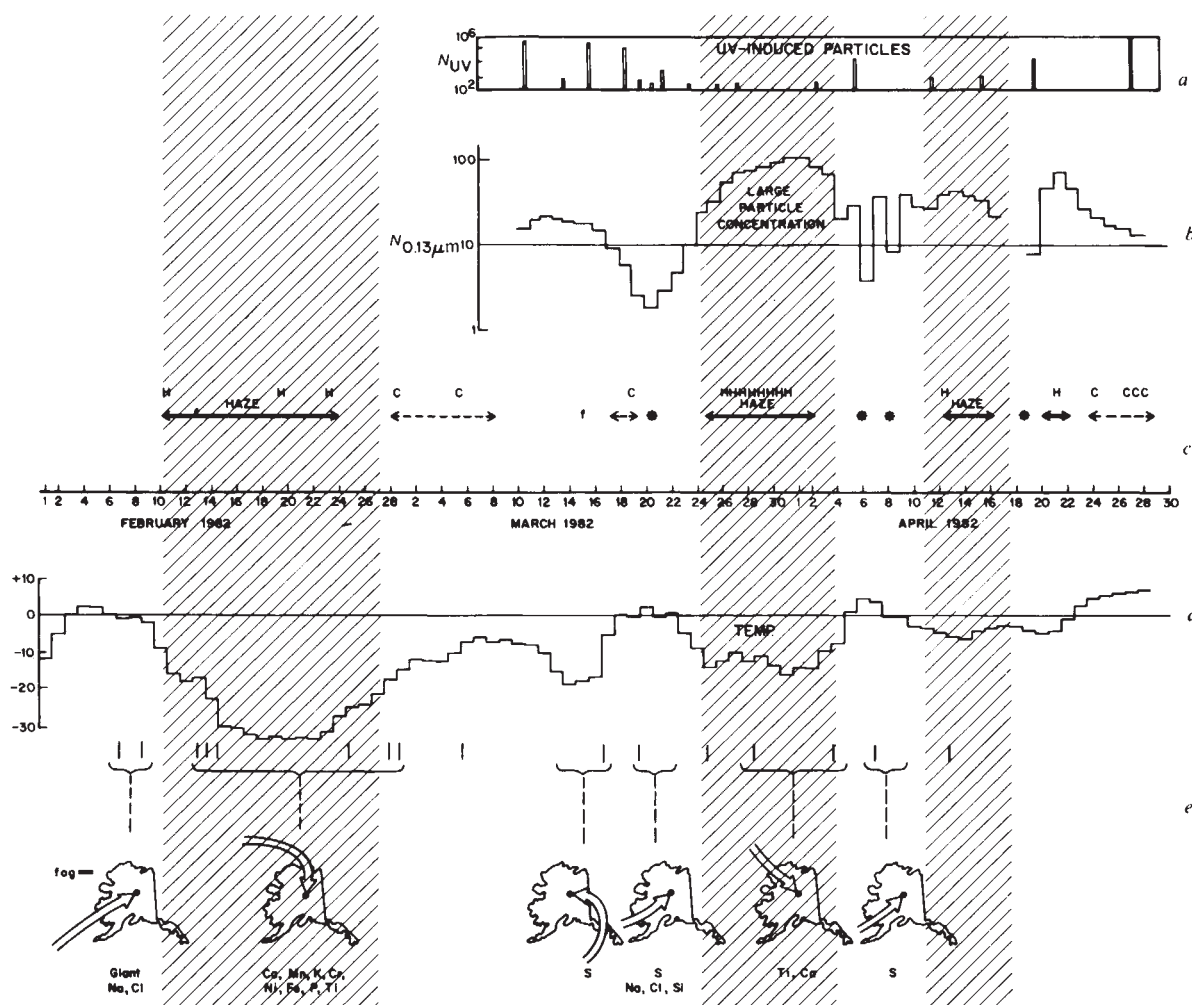


Fig. 1 Parameters associated with Arctic haze in central Alaska. Cross-hatched regions are periods when visibility degraded in central Alaska; they are associated with low air temperature and air masses entering Alaska from the Arctic, west of Alaska (arrows on the maps in *c* show incoming air trajectories). During Arctic haze episodes the concentration of particles in the size range 0.13–0.14 μm diameter (curve *b*) increased. Irradiation of the air samples with the 1,800 Å line of a low pressure Hg discharge lamp nucleated fewer particles (*a*) during Arctic air flow than during air flow from the Pacific Ocean region. H and C refer to visual observation of haze and clear, high visibility periods, respectively.

Na and Cl (marine). When air entered Alaska from the Arctic, the X-ray spectra became very complex showing the presence of various heavy metals including Ti, Cr, Mn, Fe and Ni; additionally the aerosol mass loading increased by several times.

During Arctic haze outbreaks (cross-hatched in Fig. 1) back air trajectories from central Alaska at the 500-mbar level were traced to the central Eurasian sector of the Arctic (Fig. 2). Actual transport of the aerosol material occurs probably at altitudes well below the altitude corresponding to 500 mbar, but the 500-mbar streamlines indicate the large-scale features of the transport pathway. The major synoptic feature controll-

ing the incoming air flow was anticyclonic activity over eastern Siberia and a trough over north-central Eurasia. The simultaneous occurrence of such flow conditions (which occur most frequently in the mid-winter months) and the enhanced metal X-ray signatures in aerosol suggest the presence of significant aerosol sources in central Eurasia, particularly around the area of the Taymyr Peninsula (there are no other apparent sources along the transport pathway). A Siberian source region of aerosol has also been deduced for the Canadian Arctic⁴.

Although there are various pollution sources in central

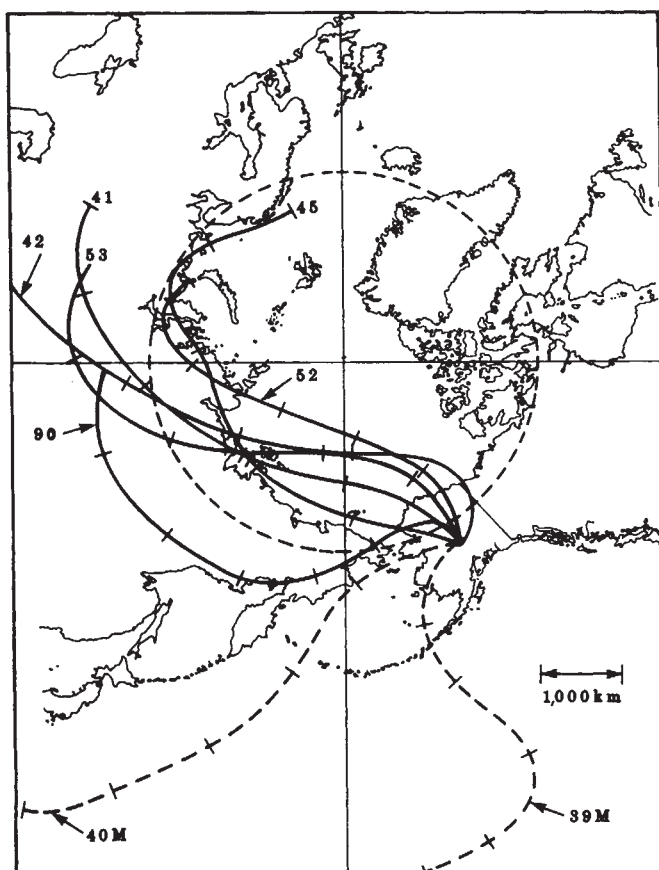


Fig. 2 Five-day back air trajectories to central Alaska at the 500-mbar pressure level. Solid lines, periods when visibility was reduced due to pervasive Arctic haze. Dashed lines, examples of trajectories during clear periods. Numbers refer to the date on which the trajectory reached central Alaska (calendar day, 1982). M indicates that marine elements (Na, Cl) were abundant in aerosol samples. The Arctic Circle is shown for reference. Tick marks on the trajectories refer to sequential 1-day intervals of travel time.

Eurasia, one that may be important is the large polymetallic ore mining-smelting complex at Norilsk (69.4°N , 88.5°E). Satellite photographs of the region (Fig. 3) show the presence of two large plumes, which can be traced for up to 40 km and which emanate from smelters processing rich sulphide Ni-Cu ores.

From the growth rate⁸ and nickel output at Norilsk in the mid-1970s⁹, the present production is estimated at 0.5 million tons of Cu and Ni annually. SO_2 emissions from roasting and reverberatory furnaces and converters at this size operation (scaled from the Sudbury Ni plant in Ontario¹⁰, which has similar ores) would be substantial, probably $\approx 10 \text{ kg SO}_2 \text{ s}^{-1}$, some portion of which would convert to submicrometre particles during transport along a pathway of several thousand kilometres¹¹. If the secondary particles have residence in the atmosphere long enough to be conserved in the travelling plume, the sulphate particle loading from Norilsk in a plume 200 km wide¹² entering Alaska is calculated to be $\sim 4 \mu\text{g m}^{-3}$ (assuming complete conversion of SO_2 , which may not occur), compared with several $\mu\text{g m}^{-3}$ sulphate mass loading actually observed in Alaska during the haze episodes. The assumption of a long residence time, say 4–7 days, is consistent with our observation (with diffusion batteries) that most particles of Arctic haze were in a mode of $\sim 0.2 \mu\text{m}$ diameter, a size too small to be efficiently removed by inertial processes and too

Fig. 3 Landsat image of Norilsk mining smelting region (14 August 1981; 0516 GMT). Distance between marks is 11 km. North is up. Two plumes from polymetallic ore refineries can be traced for 30–50 km.



large to be removed efficiently in diffusive-controlled processes¹³.

Of course, the mass loading values obtained with a simple first-order model can only be used to make order of magnitude judgements and it is not justifiable to assert that smelting activities at Norilsk are responsible for Arctic haze in Alaska. The numbers and back trajectories do suggest, however, that air emissions associated with combined industrial activity in central Eurasia may constitute a so far undetermined, but seemingly significant, portion of Arctic haze observed in the American Arctic.

We thank T. Armstrong, J. Roederer and T. Shabad for helpful comments. This research was sponsored by ONR grant N-00014-C-0435 and NSF grant DPP79-19816.

Received 7 July; accepted 24 August 1982.

1. Kerr, R. A. *Science* **205**, 290–293 (1979).
2. Kerr, R. A. *Science* **212**, 1013–1014 (1981).
3. Rahn, K. A. *Atmos. Environ.* **15**, 1447–1455 (1981).
4. Barrie, L. A., Hoff, R. M. & Daggupaty, M. *Atmos. Environ.* **15**, 1407–1419 (1981).
5. Rahn, K. A. *Atmos. Environ.* **15**, 1457–1464 (1981).
6. Armstrong, J. T. *Scanning Electron Microscopy* Vol. I, 455–466 (SEM/NC, AMF O'Hare, Illinois, 1978).
7. Carlson, T. N. *Atmos. Environ.* **15**, 1473–1478 (1982).
8. Shabad, T. & Mote, U. L. *Gateway to Siberian Resources* (Wiley, New York, 1977).
9. Central Intelligence Agency *Polar Regions Atlas* (GC 78-10040) (US Government Printing Office, Washington DC, 1979).
10. Scheider, W. A., Jeffries, D. S. & Dillon, P. J. *Atmos. Environ.* **15**, 945–956 (1981).
11. Eatough, D. J. et al. *Atmos. Environ.* **16**, 1001–1015 (1982).
12. Hage, K. D. in *Handbook of Environmental Control*, Vol. I (eds Bond, R. G., Straub, C. P. & Prober, R.) (CRC, Cleveland, 1972).
13. Twomey, S. *Atmos. Aerosols* (1977).

Female choice selects for extreme tail length in a widowbird

Malte Andersson

Department of Zoology, University of Gothenburg, PO Box 25059, S-400 31 Gothenburg, Sweden

Darwin's¹ hypothesis that male secondary sexual ornaments evolve through female preferences is theoretically plausible^{2–7}, but there is little experimental field evidence that such preferences exist^{8–10}. I have studied female choice in relation to male tail length in the long-tailed widowbird, *Euplectes progne*, and report here that males in which the tail was experimentally elongated showed higher mating success than males having normal or reduced tails. The possibility that intrasexual competition among males maintains the long tail was not supported: males with shortened tails held their territories as long as did other males. These results suggest that the extreme tail length in male long-tailed widowbirds is maintained by female mating preferences.

Male long-tailed widows have the most extreme sexual ornaments among *Euplectes*, an African genus of polygynous weaverbirds (the Ploceidae)¹¹. Reproductive adult males are black except for a red epaulet on the wing, but the most conspicuous feature is the tail: 6–8 of the 12 tail feathers are ~0.5 m long, the rest being one- to two-thirds as long. The tail during flight display is expanded vertically into a deep, long keel below the male as he flies with slow wingbeats 0.5–2 m above the territory. Displaying long-tailed widows are visible from over 1 km distance on their 0.5–3-hectare territories. The territories lie in open grassland on the Kinangop plateau, Kenya, where *E. progne* is one of the most common birds, and where the present study was performed between November 1981 and March 1982. Females are inconspicuous, being mottled brown, with short tails (~7 cm). They build their nests on the territories of the males, in the upper third of the 0.5–0.8 m-high grass, *Eleusine jaegeri*, and raise their young (2–3) unaided by the male. These features make the long-tailed widow

suitable for a test of the theory of intersexual selection of male ornaments.

Darwin¹ and Fisher² suggested that further evolution of an ornament ceases when it becomes so large that it reduces survival enough to exactly balance the mating advantage. For this to be so, females must prefer larger than normal-sized ornaments; otherwise there can be no balance between the two selection pressures. In the present experiment, females chose from males with shortened, normal or elongated tails. The Darwin–Fisher theory therefore predicts that mating success should be highest among males with elongated tails, and lowest among males with shortened tails.

The experiment included nine groups, each containing four individually colour-ringed males, of similar initial tail length and territory quality. Territory boundaries were determined by plotting on maps the locations of male displays and attacks, using cattle fences, streams and vegetation features as landmarks. In each matched group, the following treatments were randomly allocated among the four males. The tail was cut to ~14 cm in one of them; each removed feather was then attached with rapidly (~1 s) hardening cyanoacrylate glue to the corresponding feather in another male, whose tail was thus prolonged by an average of 25 cm. About 3 cm of each removed feather was first cut off and glued back on to its counterpart in the 'shortened' male, which hence was manipulated in a similar manner to the 'elongated' male. The two remaining males were controls; one was ringed only. To check whether the cut-and-glue operation influenced male behaviour or female choice, the tail of the second control male was cut off at the midpoint; each feather was then glued back on again. This operation shortened the tail by only 1 cm (~2%), which is probably not noticeable by females. Uneven joints or ends of glued feathers were trimmed with a scalpel. The joints were difficult to see from more than 1 m.

As capture and manipulation might influence subsequent behaviour, which could confound interpretation, flight displays and territorial disputes were counted in each male for 30 min 1–5 days before, and 10–14 days after the treatment.

Male mating success was estimated by the number of active nests (containing eggs or young) on each territory, for which I searched for about 1 h just after treatment of the male, and at weekly intervals for 1 month afterwards. No new clutches were laid after early January. To avoid bias, I searched each territory in proportion to its area of nesting habitat (tall, rank grass). The first count provided a standard with which to compare male mating success after tail treatment, estimated by the number of clutches laid during the remainder of the breeding season, after the day on which the male was manipulated. This use of each male as his own control reduces the importance of differences in territory quality, which influences female choice of mate in the long-tailed widow (M.A., in preparation). Counts spanned 1 month, so a few nests might represent re-laying, but this should not bias the result as treatments were randomized within each group.

Euplectes females usually seem to mate with the male on whose territory they nest^{12,13}. The possibility that a female might mate with one male and nest on the territory of another should not produce bias favouring the Darwin–Fisher prediction but should reduce the likelihood of detecting female mating preferences in the present study. Whereas attractive males in such cases receive more matings, some of their females may build on the less crowded territories of less attractive males. Group defence against nest predators is poorly developed, and there is no other evidence that female long-tailed widows would benefit from clumped breeding; the nests are well dispersed within a territory.

Females will be selected to respond to a character only if it varies among potential mates¹⁰. This is the case in long-tailed widows: the fully grown tails (no blood quills) of seven territorial males had a mean length of 49.6 cm, with c.v. = 9.4%. There was no significant correlation between male tail length and number of nests on the territory before the experiment.

However, the present randomized block design is unsuitable for detecting such a correlation; it was used to make the experiment maximally sensitive to any importance of tail length.

The two types of control males (I, cut and restored; II, only ringed—see Fig. 1) did not differ significantly in tail length, display rate or any other measured character. Therefore, they should be equivalent from a female perspective, so I have treated them as one category below, with two representatives in each group of four males.

Before tail treatments, there were only minor differences in mating success between 'shortened', control and 'elongated' males (Fig. 1a). After treatment, however, mating success changed as predicted by the female choice hypothesis, with lowest success for males having shortened tails, and highest

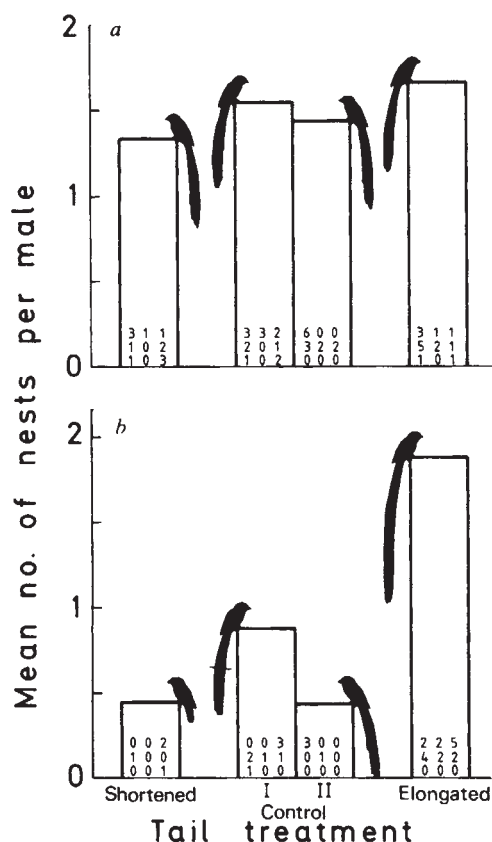


Fig. 1 Mating success in male long-tailed widows subjected to different tail treatments. *a*, Mean number of active nests per territory for the nine males of the four treatment categories, before the experiment. Numbers of nests for the nine males in each of the four categories are given at the bottom of the bars, always in the same order of matched 4-group. *b*, Number of new active nests in each territory after treatment of males. The following procedure was used to test for differences in mating success between the treatment categories. For each male I calculated the difference between the number of new active nests on his territory during the remainder of the breeding season after tail treatment, and the original number of nests on the territory before tail treatment. (Subtraction of the original number of nests reduces the influence of initial variation among males and territories.) These differences were used for matched comparisons (with respect to each group of four males) between shortened, control and elongated males. As predicted by the Darwin-Fisher theory of sexual selection, males with elongated tails became significantly more successful (as measured above, compared with before the tail treatment) than shortened males and control males II ($P < 0.05$ in both cases, paired randomization tests). For the four categories tested together, there was a significant trend of increasing mating success as tail length increased from shortened via control to elongated males ($P = 0.03$, Pitman randomization test¹⁸ adapted to the present experimental design; no standard test fits this situation, with two different treatments (shortened and prolonged tails) and two control categories, and an alternative hypothesis which predicts a specified trend).

success for males with elongated tails (Fig. 1b). Hence, tail length apparently did influence mate choice: females preferred those males having the longest tails. As the main difference was between 'elongated' males and other males (including controls), the difference did not result from possible destruction of species-specific features in males having shortened tails.

Another possible explanation is that 'shortened' males became less active in their courtship behaviour, or that males with elongated tails became more active. However, the only indication of a difference was the reverse of this (but not significant: $P > 0.95$, Friedman two-way analysis of variance). Males having reduced tails increased their average rate of flight display slightly (from 9.2 to 10.3 displays per 30 min), whereas control males showed a decrease (from 10.3 to 6.9), as did males with elongated tails (from 10.8 to 7.8). Therefore, changes in male behaviour were probably not responsible for the higher success of elongated males.

In choosing her mate, a female should respond to the quality of his territory, on which she nests^{10,14}. However, due to the randomization within each group of four males and the use of each male as his own control, territory differences cannot explain the higher mating success of males with elongated tails. As behavioural differences were also excluded, I conclude that the changes in male tail length caused the differences in the attraction for females.

As is implicit in the Darwin-Fisher theory of sexual selection, females preferred males having tails that were longer than normal. This is expected if females are attracted by 'supernormal stimuli'^{4,6,8}. Such a preference can evolve if asymmetrical selection shapes female responsiveness^{4,15}. One possibility is that a male's sexual ornaments reflect his overall phenotypic and genotypic quality, so that females choosing highly ornamented males bear offspring having high expected fitness⁹. However, it is unknown whether fitness in nature is heritable enough to influence female choice of mate.

Highly adorned males can be favoured by active mate choice, where females compare males before accepting one, but also by easier detection¹⁶. This latter advantage may have contributed to the evolution of the long tail and the flight display in the long-tailed widow. The lateral surface of the displaying male is enlarged 2–3 times by the tail, making him correspondingly easier to discover from a distance in the open habitat. However, neighbouring males often display simultaneously, and females sometimes visit several males in rapid succession with ample opportunity for comparisons. The long tail is therefore probably maintained at least partly through active female discrimination among males.

Alternatively, ornaments may be favoured by intrasexual selection among males competing for territories or hierarchy ranks^{9,10,17}. This hypothesis predicts that males having shortened tails are least efficient at holding a territory, and that males with elongated tails are most efficient; this was not supported. Most males remained on their territories until February, when the nesting was over, and territory tenacity did not differ among treatment categories ($P > 0.6$, Pitman randomization test). There was no evidence of increased territory size in males having elongated tails. Males with shortened tails took off and defended their territories more often than other males but the difference was not significant ($P > 0.1$, Friedman two-way analysis of variance); this may indicate more intrusions on their territories. However, 'shortened' males also increased their rate of flight display (see above), usually performed when the male is alone on the territory, or is visited by females. Easier flight in these males relieved of the unwieldy tail, which is carried only during the breeding season, may explain their higher rate of territory defence as well as flight display (insofar as the non-significant differences are real).

The results presented here support Darwin's¹ hypothesis that certain male ornaments are favoured by female mate choice, and probably evolved through it.

I thank the Office of the President, Nairobi, for permission to do field work in Kenya; Derek Pomeroy for assistance with

the study; Kuria Mwaniki and Uno Unger for help in the field, Björn Rosander for statistical advice; and Conny Askenmo, John Maynard Smith, and Peter O'Donald for comments on the manuscript. The study was supported by the Swedish Natural Sciences Research Council.

Received 21 June; accepted 2 September 1982.

1. Darwin, C. *The Descent of Man, and Selection in Relation to Sex* (John Murray, London, 1871).
2. Fisher, R. A. *The Genetical Theory of Natural Selection* (Clarendon, Oxford, 1930).
3. Maynard Smith, J. in *A Century of Darwin* (ed. Barnett, S. A.) 231–245 (Heinemann, London, 1958).
4. O'Donald, P. *Theor. Populat. Biol.* **12**, 298–334 (1977).
5. O'Donald, P. *Genetic Models of Sexual Selection* (Cambridge University Press, 1980).
6. Lande, R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3721–3725 (1981).
7. Kirkpatrick, M. *Evolution* **36**, 1–12 (1982).
8. Halliday, T. R. in *Behavioural Ecology, An Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 180–213 (Blackwell, Oxford, 1978).
9. Andersson, M. *Biol. J. Linn. Soc.* **17**, 375–393 (1982).
10. Searcy, W. A. *Rev. Ecol. Syst.* **13** (in the press).
11. Craig, A. J. F. K. *J. Orn.* **121**, 144–161 (1980).
12. Lack, D. *Ibis* **77**, 817–836 (1935).
13. Craig, A. J. F. K. *Ostrich* **45**, 149–160 (1974).
14. Orians, G. H. *Am. Nat.* **103**, 589–603 (1969).
15. Staddon, J. E. R. *Am. Nat.* **109**, 541–545 (1975).
16. Parker, G. A. in *Current Problems in Sociobiology* (Cambridge University Press, in the press).
17. Borgia, G. in *Sexual Selection and Reproductive Competition in Insects* (eds Blum, M. S. & Blum, N. A.) 19–80 (Academic, New York, 1979).
18. Bradley, J. V. *Distribution-free Statistical Tests* (Prentice-Hall, Englewood Cliffs, 1968).

Self-pituitary grafts are not rejected by frogs deprived of their pituitary anlagen as embryos

Louise A. Rollins-Smith & Nicholas Cohen

Division of Immunology, Department of Microbiology, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, USA

In the present study, we have adopted the model of Triplett¹ to reinvestigate the timing of development of immunological tolerance to self-organ-specific antigens. We have removed pituitary or eye² anlagen from frog embryos before development of the immune system and returned them at a later time as differentiated organ implants to their now immunocompetent larval or adult original owners. If immunological tolerance to these putative organ-specific self-antigens occurs at an early and fixed time period, then organ-deprived hosts, lacking the opportunity to become tolerant, would be expected to reject such implants¹. Our results show that self-implants were never rejected whereas control allogeneic implants were usually rejected by larval hosts and were always rejected by adult hosts. These data, which contrast with those reported by Triplett, suggest that frogs, and perhaps other higher vertebrates, can become tolerant to self-organ-specific antigens throughout life.

We recently reported² experiments in which eye anlagen were removed from *Rana pipiens* and *Xenopus laevis* (*X. laevis* and *X. laevis-gilli* hybrid clones³) embryos. When the enucleated embryos developed into immunocompetent larvae, they were implanted with either their own (previously parked) or an isogeneic (cloned) differentiated eye. All self-grafts in intact hosts or enucleated hosts survived in almost perfect condition for as long as they were observed (>100 days; Fig. 3a, b). In contrast, allogeneic eyes were rejected by about half of the intact larval *Xenopus* and *Rana* hosts.

Because only some of the larval hosts rejected allogeneic eyes², and because other studies had demonstrated a greater degree of tolerance of skin allografts in larval than adult hosts⁴, we enucleated cloned (LG15) *Xenopus* embryos (Nieuwkoop and Faber⁵ stages 26–32) and implanted them with an isogeneic eye 2 months after they metamorphosed. In such postmetamorphic *Xenopus*, all allogeneic eye implants were rejected vigorously and rapidly (most within 20 days), whereas isogeneic implants on intact or embryonically enucleated hosts survived for as long as they were observed (2–24 months, Figs 1, 3c). Examination of serially sectioned isogeneic implants fixed at

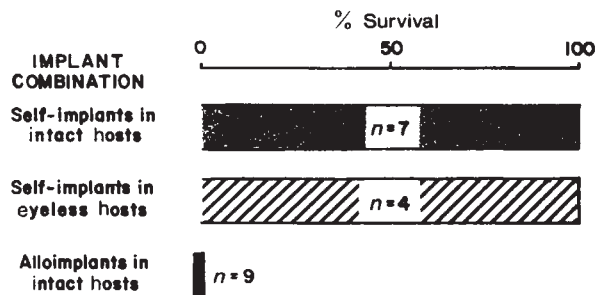


Fig. 1 Per cent survival of self- (▨) and allogeneic (■) eye implants in intact postmetamorphic *Xenopus* and of self-implants in embryonically enucleated postmetamorphic *Xenopus* (▨) at 90 days post-implantation. The number of frogs in each group is indicated within the bars. All enucleations were performed during stage 26–32. One animal, however, developed a partial regenerated eye that was re-extirpated at about stage 43.

1–2 yr post-implantation revealed no lymphocytic infiltration or other evidence of immune destruction.

To minimize the criticism that the eye may not be a good model organ to study development of unresponsiveness to organ-specific antigens and to determine the generality of our observations with the eye, we used the pituitary as a test organ. Triplett reported that most (10/13) self-pituitary implants were rejected by their embryonically hypophysectomized larval tree frog hosts with a mean graft survival time of ~40 days. This study suggested that the pituitary has organ-specific transplantation antigens, and provided experimental support for the idea that unresponsiveness to these antigens must develop in a fixed early time period relative to lymphocyte ontogeny. In our studies, pituitary anlagen were extirpated from Shumway⁶ stage 17–18 *R. pipiens* embryos and 'parked' orthotopically on previously hypophysectomized stage 17–18 sibling hosts. After 50–60 days, we dissected the now differentiated pituitaries free from the brain of the 'parking' host and implanted them in the dorsal tail-fins of the original hosts. The hypophysectomized hosts were pale and their growth was retarded (Fig. 3d). After implantation of either a self- or an allogeneic pituitary, they began to darken due to melanophore stimulating hormone (MSH) release by the implanted gland, and their growth rate increased. Six of nine allogeneic pituitaries were rejected within 20 days. Recipients of these implants again became pale. In marked contrast, all recipients of the self-pituitaries remained dark for as long as they were observed (three for >90 days, Figs 2, 3e). Examination of serially sectioned heads of the hypophysectomized hosts showed no identifiable anterior pituitary cells in association with the brain of any of the six larvae examined (compare Fig. 3f and g). Thus, the darkening of pituitary implanted animals seemed to result from MSH production by the self-implant rather than from any anterior pituitary tissue remaining after embryonic ablation.

Three technical differences between our pituitary grafting experiments and those of Triplett may bear on our discordant observations. First, our experiments were done with *R. pipiens* while his were done with *Hyla regilla*. Thus, although the anlagen were extirpated at comparable stages of embryonic development, and the differentiated organs were reimplanted after about the same time lag, we cannot exclude the possibility that there are species-specific differences in the maturation of immunity and a narrower period in *Hyla* in which self-tolerance may occur.

A second experimental difference was the site in which the pituitary anlagen differentiated. In our experiments, the anlagen were 'parked' orthotopically in previously hypophysectomized age-matched sibling embryos. In contrast, Triplett 'parked' the anlagen under the ectoderm in the tail region in rather older nonsibling larvae¹. In this site, the graft induced formation of a heavily pigmented host-derived fibrous connective tissue capsule that was always transplanted together with the pituitary¹.

Our implants were free of host-derived connective tissue. Whether this allogeneic contaminant (an estimated 20% of the mass of the transplant; E. L. Triplett, personal communication) contributed to the destruction of the self-pituitary graft, is conjectural. Triplett, in fact, attempted to control for allocontamination in a set of experiments in which he removed, 'parked' and then returned a partial gland. Although he observed that all seven partial self-grafts survived, he also reported the survival of four of six allogeneic pituitaries in partially hypophysectomized recipients. This finding, which is consistent with the high frequency of allotolerance observed in larval *Xenopus*⁴ and *Rana*², makes the data from the partial-hypophysectomy experiment difficult to interpret.

A third technical difference is that Triplett supplemented the diet of completely hypophysectomized animals with beef thyroid powder "about 60 days" after hypophysectomy. Our hypophysectomized frogs never received exogenous hormone and did not progress beyond Taylor-Kollros⁷ stage III until after they had received a pituitary implant. That exogenous thyroxine treatment might reveal a net self-cytotoxic rather than tolerogenic response in *Hyla* cannot be excluded. However, it must be pointed out that in our experiments with eye implants, enucleated frogs had an intact endocrine system and self-cytotoxicity was not the net result following implantation of a self-eye. Further, at least alloreactivity of hypophysectomized frogs can develop without addition of exogenous thyroxine, as more than half of the hypophysectomized recipients of allogeneic pituitary grafts in our experiments were capable of rejecting those grafts. Perhaps a better understanding of the relationships between the developing endocrine and immune systems in amphibians might shed light on our discordant data.

Our failure to confirm Triplett's observation may mean that tolerance to organ-specific antigens can occur throughout larval life and, at least in the case of the eye, during adult life. Alternatively, because it was Triplett's original observation of self-rejection that suggested the existence of pituitary-specific antigens, our failure to confirm his observation of rejection may mean either that such antigens do not exist in *R. pipiens* or that they are of such a distribution or immunogenicity as to be incapable of evoking a net destructive response. Reports of eye-specific antigens in other species⁸ and of autoimmune responses directed against the mammalian pituitary^{9,10} do provide independent evidence of pituitary- and eye-specific antigens; a rigorous demonstration of their existence in amphibians awaits further study.

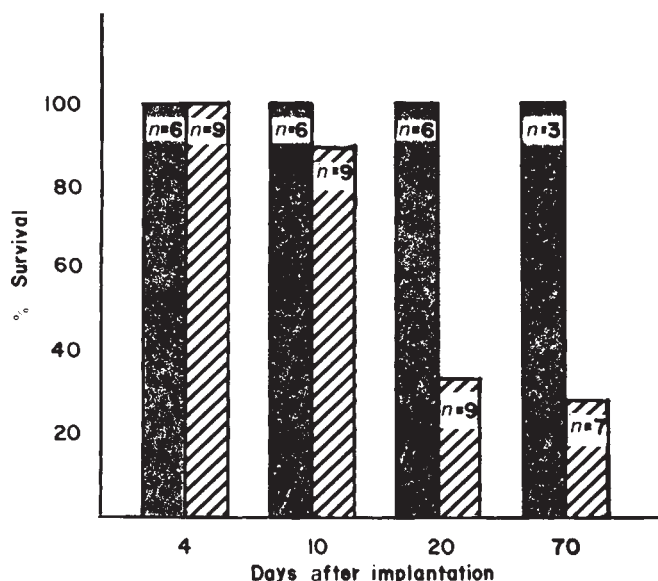


Fig. 2 Per cent survival of self- (■) and allogeneic (▨) pituitary implants in embryonically hypophysectomized *Rana* larvae. The number of larvae in each group is indicated within the bars.

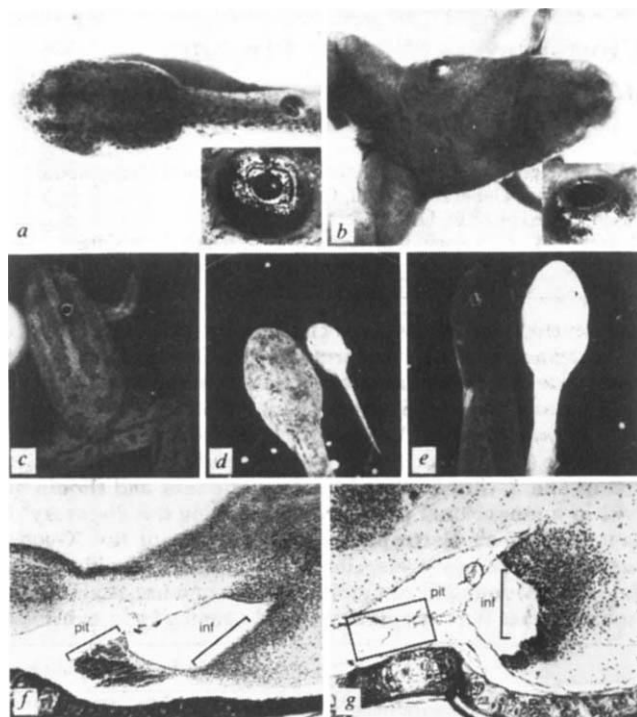


Fig. 3 *a*, *Rana pipiens* tadpole, enucleated as an embryo and grafted during larval life with a self-eye. Photographed 75 days after implantation ($\times 1.9$), insert ($\times 8.9$) (modified from ref. 2). *b*, Postmetamorphic *Xenopus laevis*, enucleated as an embryo, with self-eye implant placed during larval life. Photograph taken about 1 yr after implantation ($\times 1.4$), insert ($\times 6.2$) (from ref. 2). *c*, Postmetamorphic *Xenopus laevis-gilli*, enucleated as an embryo, with isogeneic eye implanted 2 months after metamorphosis. Photograph taken about 1½ yr after implantation ($\times 1.4$). *d*, *Rana pipiens* tadpoles, 41 days after they had been hypophysectomized as embryos. Tadpole on the left was hypophysectomized and served as the 'parking host' for the pituitary of the tadpole on the right. Note that the 'parking host' is of normal size and pigmentation while the hypophysectomized tadpole is small and pale ($\times 1.4$). *e*, *Rana pipiens* tadpoles. Tadpole on left was hypophysectomized as an embryo, implanted with its self-pituitary 60 days later, and photographed about 60 days after implantation. This tadpole survived with an intact implant 91 days post-implantation. Note the abnormally dark pigmentation due to unregulated MSH production by the implanted gland. The tadpole on the right shows normal pigmentation of a tadpole about 3 months of age. The animal appears to be blanched because the photograph was underexposed to reveal the detail of the darkly pigmented frog with the self-pituitary implant ($\times 1.0$). *f*, Sagittal section through the brain of a normal tadpole (stage IX) indicating the infundibulum (inf) of the diencephalon and the pituitary (pit) ($\times 66.8$). *g*, Sagittal section through the brain of an embryonically hypophysectomized tadpole that received a self-pituitary implant (tadpole pictured on left in *e*). The truncated infundibulum (inf) is indicated and the empty black rectangle indicates the approximate location of the pituitary (pit) had it not been removed ($\times 66.8$).

This research was supported by USPHS grants F32 CA06375, HD 07901 and AI 18319. We acknowledge the animal care given by Deborah Sussman, Eleanor Berger and Nancy Efrige.

Received 19 March; accepted 1 September 1982.

1. Triplett, E. L. *J. Immun.* **89**, 505-510 (1962).
2. Rollins, L. A. & Cohen, N. in *Development and Differentiation of Vertebrate Lymphocytes* (ed. Horton, J.) 203-214 (Elsevier, Amsterdam, 1980).
3. Kobel, H. R. & Du Pasquier, L. *Immunogenetics* **2**, 87-91 (1975).
4. DiMarzio, S. & Cohen, N. *Immunology* **45**, 39-48 (1982).
5. Nieuwkoop, P. D. & Faber, J. *Normal Table of Xenopus laevis* (Daudin), 1-252 (North-Holland, Amsterdam, 1967).
6. Shumway, W. *Anat. Rec.* **78**, 139-147 (1940).
7. Taylor, A. C. & Kollros, J. *J. Anat. Rec.* **94**, 7-23 (1946).
8. Rahi, A. H. S. & Garner, A. *Immunopathology of the Eye*, 1-343 (Blackwell Scientific, London, 1976).
9. Asa, S. L., Bilbao, J. M., Kovacs, K., Josse, R. G. & Kreines, K. *Ann. intern. Med.* **95**, 166-171 (1981).
10. Topliss, D. J. & Volpe, R. *Ann. intern. Med.* **95**, 227-229 (1981).

Clonal interaction in tumours

M. F. A. Woodruff*, J. D. Ansell, G. M. Forbes*,
J. C. Gordon*, D. I. Burton & H. S. Micklem

*Medical Research Council Clinical and Population Cytogenetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK
Department of Zoology, University of Edinburgh, The Kings Buildings, West Mains Road, Edinburgh EH9 3JT, UK

The development of cancer is contingent on the emergence of at least one clone of transformed cells. One method used to investigate whether human tumours are monoclonal depends on the mosaicism in the normal tissues of women heterozygous for the two forms of the enzyme glucose-6-phosphate dehydrogenase (G-6-PD)¹. This mosaicism results from the inactivation of one X chromosome in all somatic cells and should not exist in a monoclonal population. Following the discovery² in feral mice of an electrophoretic variant (A) of the X-coded enzyme phosphoglycerate kinase (PGK-1) which differs from the form (B) found in common laboratory mouse strains it was reported³ that fibrosarcomas induced chemically in hybrids of

feral and laboratory-bred mice expressed both enzyme phenotypes, but the conclusion that both were expressed by neoplastic cells was based solely on morphological evidence. The development of histocompatible substrains of mice homozygous for one or other alloenzyme has made it possible to study the clonal composition of tumours under experimental conditions in which the neoplastic status of subpopulations of cells can be verified by transplantation. The experiments we now report, while confirming that murine fibrosarcomas are often pleoclonal, show that the clonal composition may change markedly during tissue culture and on transplantation to congenic hosts. These changes presumably reflect changes in the growth kinetics of differentiating subpopulations of the tumour. Cloned sublines are less readily transplantable than uncloned tumour cell populations, and some sublines are less readily transplantable than others; this suggests that sublines resistant to a host's attack are selected on transplantation or that some sublines require the cooperation of others to survive. We postulate that changes in clonal composition occur also during tumour development, metastasis and recurrence.

Fibrosarcomas were induced in adult female CBA backcross mice heterozygous at the PGK locus⁴ (*Pgk-1^a/Pgk-1^b*, abbreviated to AB) by subcutaneous (s.c.) injection of methylcholanthrene (MC) dissolved in tricaprillin (S tumours) or s.c.

Table 1 PGK-1 alloenzyme assays of uncloned cultures

No. of tumours	Reference numbers of tumours	Results	Interpretation (types of neoplastic clones)
19	S1, 6, 7, 8, 9, 11, 15; D6, 7, 9, 12, 16, 17 D1, 5, 8, 18 S3; D2	A in all cultures; B in none A in all cultures, B trace to 10% in some primary cultures A in all cultures; B trace to 10% in some primary and 2nd generation cultures	A only
1	S5	B in all cultures; A 50% falling to 20% in primary cultures only	B only
2	S12, 14	A in all cultures; B increased from trace in primary to 20% in 3rd generation cultures, then disappeared	A only or A and B
2	D4, 14	B in all cultures; A increased from trace in early primary to 20% in 3rd gen. cultures, then disappeared	B only or A and B
10	S13; D15 S4 S2; D10 D13 D3 S10 D11 D19	A in all cultures; B 30–50% in primary and 2nd generation cultures, then declined and disappeared A in all cultures; B 80% in primary culture, then declined and disappeared A in all cultures, B 50% in early primary cultures, persisted to 4th to 6th culture generation in amounts which varied from trace to 50% and then disappeared A with trace B in primary and early 2nd generation cultures; B only, or with trace to 20% A in all subsequent cultures (to 8th generation) B in all cultures; A increased during primary cultures from trace to 50%, persisted to 5th culture generation in amounts which varied from 10–60% and then disappeared B in all cultures; A absent in primary but increased in 2nd generation to 50% and persisted in all subsequent cultures in amounts which varied from 20–50% A and B in substantial amounts in all cultures, A being the larger component in some and B in others A and B in equal amounts in primary culture; subcultures did not grow	A and B

Fibrosarcomas were induced by s.c. injection of 0.5 mg methylcholanthrene (MC) dissolved in tricaprillin (S tumours) or s.c. implantation of a disk (6 mm diameter) of Millipore membrane (0.22 µm pore diam.) impregnated with 0.1 mg MC (D tumours). Sarcomas developed within 9 months in 35 out of 38 mice and were numbered in accordance with the time they took to develop. Cell suspensions were prepared with Dispase⁶ from tumours collected 113–211 days after carcinogen administration, washed, and resuspended in Ham's F10 medium with 10% fetal calf serum (FCS). As a rule the hosts were killed, but in a few animals the tumour-bearing hind limb was removed surgically and in two of these the tumour recurred at the site of amputation. Tissue culture flasks (Falcon 75 cm²) were seeded with 10⁷ viable cells and incubated at 37 °C in an atmosphere containing 5% CO₂. After culture for 18 h to 20 days, non-adherent cells were discarded and adherent tumour cells detached by brief exposure to trypsin (0.07%) and EDTA (0.027%) were used to set up subcultures. This procedure eliminates nearly all the leucocytes (which are non-adherent) and macrophages (which are strongly adherent), but not fibroblasts. Samples from primary cultures and subcultures, and usually also of suspensions prepared from the whole tumour and from host liver, were frozen and thawed in lytic buffer, and the approximate proportions of the two alloenzymes were determined by gel electrophoresis, using a modification of the linked enzyme assay developed by Buchet *et al.*⁸, the production of NADPH being visualized by the reduction of a tetrazolium dye, thiazolyl blue, to its formazan derivative.

implantation of a disk of Millipore membrane impregnated with MC⁵ (D tumours). The relative amounts of the two alloenzymes were determined in whole tumour suspensions⁶ and uncloned cultures after gel electrophoresis (Table 1). As expected, both A and B components were found more often in whole tumour suspensions (17 out of 24 studied) than in primary cultures of the same tumour (11 out of 24), which should contain fewer normal cells. With many tumours the A/B ratio fluctuated greatly during culture. Contrary to the earlier finding of Reddy and Fialkow⁷, there was a marked preponderance of tumours with an A component; this is not surprising, however, because the *Pgk-1* locus is closely linked to another locus, *Xce*, which determines the probability of the X-chromosome being inactivated⁷; the resulting ratio of A cells to B cells in the normal tissues of AB mice of our CBA stock is about 70:30.

Clones were isolated from 10 tumours, and samples from cloned and uncloned cultures were injected into homozygous females (*Pgk-1^b/Pgk-1^b*, abbreviated to BB) or hemizygous males (*Pgk-1^a/Y*, abbreviated to AY); when tumours developed these were assayed in the same way as the primary tumours (Table 2). Other samples were used for karyotyping and determination of the cellular DNA content by automated scanning⁹ of Feulgen-stained preparations (1,500–4,000 cells).

With four tumours (D11, D13, D15, S10), we have formal proof of the existence of both A and B neoplastic components, since on transplantation to BB mice each gave rise to some tumours with only an A component and others with no detectable A component. In the case of D11 and S10, many A clones and B clones have been isolated and shown to be polyploid. Two A clones from D11 (Nos 6 and 10) had a modal chromosome number of 98, and automated scanning of one of these showed a mean DNA content 2.6 times that of normal diploid cells; one B clone (no. 12) had a modal chromosome number of 79 and mean DNA content 1.9 times that of normal cells. Clones 6 and 10 gave rise to A tumours, and clone 12 to a B tumour, when transplanted to BB mice. With tumour D15 the B neoplastic component disappeared in the course of tissue culture. Tumour S10 is unusual in that we have isolated A clones, B clones, and 'clones' expressing both A and B. Six of the latter were harvested from 192 wells seeded with, on average, 1 cell; and since no colonies developed in 139 wells the mean number of clonogenic cells seeded (assuming a Poisson distribution) was probably about 0.3 per well. It seemed just possible that these supposed clones were mixtures, but on recloning at the same and more extreme dilution we have obtained 19 'subclones' which still express both A and B, 20 which express only B and none which express only A. While the possibility of mixed populations of cells growing in symbiosis is still not formally excluded, it seems likely that AB clones have arisen from hybrid cells formed by fusion, either in the mouse or in tissue culture, of an A cell and a B cell, at least one of which had undergone transformation, and that disappearance of the A component in some cases on recloning was due to chromosome loss. An alternative possibility is that there has been reactivation of the inactive X chromosome in a cell which prior to transformation expressed only A or B. Preliminary observations based on flow cytometry of cells stained with the dye Hoechst 33342 have shown a higher proportion of diploid cells in suspensions from the original tumour and primary cultures than from later cultures. This may be due entirely to selection, but raises the question of whether some of the tetraploid and hypertetraploid clones we have found in other tumours may also have developed from hybrid cells formed during tissue culture. Karyotypic analysis and studies with new tumours may resolve some of these uncertainties.

Clones from five tumours (D11, D13, D14, S2, S10) which transplanted readily in the form of uncloned cultures were transplanted to normal BB mice, and in some cases also to thymectomized, irradiated BB mice protected with cytosine arabinoside¹⁰, by s.c. injection of 1.5×10^6 viable cells. Only 4 out of 21 clones grew in the normal mice, whereas 6 out of 6 which had failed to grow in the normal mice grew in the

Table 2 Characterization of clones, transplants and recurrent tumours

No. of tumour	No. and type of clones isolated	Transplants		
		Material transplanted	Neoplastic components proved	Recurrence of primary tumour
D17	6A			
S3	12A			
S5		Primary tumour	B*	B
D4	60B	Primary tumour	B*†	
D14	16B	Primary tumour	B*	
		B clone	B*†	
S13				A
D15		Primary tumour	A†B*	
		1st generation transplant	A*†	
		2nd generation transplant	A†	
S4		Primary tumour	A†	
		1st generation transplant	A*	
		2nd generation transplant	A*†	
		A clone	A†	
S2	20A‡			
D10	15A			
D13	1A, 6B	Unc cloned culture	B*†	
		A clone	A†	
D3	78B	Primary tumour	B*	
S10	A, B, AB§	A clone	A†	
		B clone	B*	
		AB clone	A†B*†	
D11	21A, 36B	Primary tumour	B*	
		A clone	A†	
		A clone	A†	
		B clone	B*	

Clones were isolated either in wells of microtest plates (Falcon 30401) seeded with 0.2 ml of a suspension containing tumour cells and irradiated (60 Gy) human fibroblasts which yielded on average 1, 2 or 5 tumour cells and 2,000 fibroblasts per well, or by ring cloning in small (25 cm²) flasks.

* Proof based on development of tumours with A component *only* in AY mice or B component *only* in BB mice.

† Proof based on development of tumours with A component in BB mice or B component in AY mice.

‡ One apparent clone of type B was isolated but on transplantation this yielded only A tumours in AY mice. On retesting, the 'clone' was found to contain a trace of A in addition to the main B component.

§ See text.

|| These clones grew only in, or after passage in, thymectomized, irradiated mice.

thymectomized mice. After passage in thymectomized mice, 5 of the 6 tumours grew when retransplanted to normal mice. These preliminary findings suggest that clones with a high capacity for survival are selected when pleoclonal populations are transplanted; in addition, some clones may require the cooperation of others to survive. The successful transplantation of five clones to normal mice after passage in thymectomized, irradiated mice suggests further that cells resistant to surveillance may emerge when tumours grow under conditions in which surveillance is impaired. The features which favour survival of a clone have not been elucidated, but possibilities to be considered include lack of strong rejection-inducing antigens; failure, despite the presence of such antigens, to evoke an effective immune response; and relative insensitivity to surveillance mediated by NK cells and macrophages.

Our results highlight the problems involved in assessing the clonality of tumours with X-linked markers, whether in man or mouse, and lend support to the view that tumours of pleoclonal origin may appear, at a particular time, to be monoclonal because all clones except one have been competitively eliminated or reduced to the point of being undetectable.

It seems likely^{12,13} that clonal interaction plays a significant role in the life history of pleoclonal tumours in general. At

present we do not have sufficient data to construct models of the cell population kinetics of such tumours, but it is possible to suggest some general principles on which these might be based: (1) The number of clones present at any given time will depend on the number of cells originally transformed and the growth rate of each clone. The number of cells transformed will depend *inter alia* on the nature and dose of the carcinogen. The growth rate of each clone may be influenced by inherent properties of the transformed cells, the host reaction, and interaction between the clone in question and other clones of transformed or initiated cells. (2) If growth rates remain constant the clone with the greatest growth rate will eventually outgrow all others, but this does not necessarily exclude the persistence of small numbers of very slowly growing or dormant cells. (3) Growth rates may change suddenly, and dormant cells may resume cycling. Such changes may be influenced by treatment but may also occur for no apparent reason. Clones which grow slowly or remain dormant in primary tumours may grow rapidly in metastases or locally recurrent tumours. (4) The symbiotic relationship between different clones may be antagonistic, neutral or mutualistic.

We thank Drs James Tucker, Daryll Green and John Hay, respectively, for undertaking automated scanning and flow cytometry, and preparing thymectomized irradiated mice. M.F.A.W., G.M.F. and J.C.G. are indebted to Professor H. J. Evans for the privilege of working in his unit, and to the MRC for a project grant. H.S.M. and J.D.A. thank the Melville Trust for a grant. D.I.B. is supported by an MRC studentship.

Received 1 June; accepted 3 August 1982.

1. Fialkow, P. J. *Adv. Cancer Res.* **15**, 191-226 (1972).
2. Nielsen, J. T. & Chapman, V. M. *Genetics* **87**, 319-327 (1977).
3. Reddy, A. L. & Fialkow, P. J. *J. exp. Med.* **150**, 878-887 (1979).
4. Burton, D. L., Ansell, J. D., Gray, R. A. & Micklem, H. S. *Nature* **298**, 562-563.
5. Woodruff, M. F. A., Bard, J., Ross, A. & Forbes, G. M. *Proc. R. Soc. B* **211**, 269-286 (1981).
6. Matsumura, T., Yamanaka, T., Hashizume, S., Irie, Y. & Nitta, K. *Jap. J. exp. Med.* **45**, 377-382 (1975).
7. Johnston, P. G. & Cattanch, B. M. *Genet. Res.* **37**, 151-160 (1981).
8. Bücher, T., Bender, W., Fundele, R., Hofner, H. & Linke, I. *FEBS Lett.* **115**, 319-324 (1980).
9. Tucker, J. H. *J. Histochem. Cytochem.* **27**, 613-620 (1979).
10. Steel, G. G., Courtenay, V. D. & Rostom, A. Y. *Br. J. Cancer* **37**, 224-230 (1978).
11. Woodruff, M. F. A. *The Interaction of Cancer and Host—Its Therapeutic Significance*, 123-133 (Grune and Stratton, New York, 1980).
12. Poste, G., Doll, J. & Fidler, I. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6226-6230 (1981).
13. Woodruff, M. F. A. *Br. J. Cancer* **46**, 313-322.

Dihydroouabain is an antagonist of ouabain inotropic action

T. Godfraind, J. Ghysel-Burton & A. De Pover

Laboratoire de Pharmacodynamie Générale et de Pharmacologie, Université Catholique de Louvain, UCL 7350, Avenue Emmanuel Mounier, 73, B-1200 Bruxelles, Belgium

The Na^+ , K^+ -pump controls a wide variety of cellular systems and its inhibition by cardiac glycosides modifies important physiological functions and evokes several pharmacological effects (refs 1, 2 and refs therein). However, not all the actions of cardiac glycosides can be attributed to Na^+ , K^+ -pump inhibition and several observations show that, at low doses, cardiac glycosides stimulate the pump³⁻⁵. It has been proposed that their positive inotropic effect could be the sum of two processes: the inhibition of the pump and a still unknown additional inotropic mechanism⁶. In guinea pig heart, low doses of ouabain interact with high-affinity binding sites, which differ from the lower-affinity sites responsible for Na^+ , K^+ -pump inhibition^{3,7-9}. It has been suggested that ouabain interaction with these high-affinity sites could be responsible for the additional inotropic

mechanism⁶. The existence of two classes of ouabain-binding sites has been documented not only in guinea pig heart, but also in dog¹⁰, rat^{11,12} and human heart¹³. Dihydroouabain, a derivative of ouabain in which the lactone ring is saturated, is about 50-fold less potent than ouabain as an inhibitor of Na^+ , K^+ -pump and does not stimulate the pump at low doses⁷. Its inotropic effect can be entirely accounted for by the inhibition of the pump^{6,14}. We have examined the pharmacological action of ouabain in the presence of dihydroouabain and report here that dihydroouabain reduces ouabain inotropic action but not Na^+ , K^+ -pump inhibition.

Figure 1 shows the inhibition of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ activity (Fig. 1a) and of specific ^3H -ouabain binding (Fig. 1b) in guinea pig heart microsomes as a function of ouabain or dihydroouabain concentration. With both glycosides, the slope of the Hill plot was equal to 1 for $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ inhibition. In contrast, the slope was 0.64 ± 0.001 for the displacement of ^3H -ouabain. This confirms the previously reported heterogeneity of ouabain binding sites^{4,15} and shows that dihydroouabain competed with ^3H -ouabain at both high- and low-affinity sites. In view of this finding, we have examined whether dihydroouabain could interact with the action of ouabain.

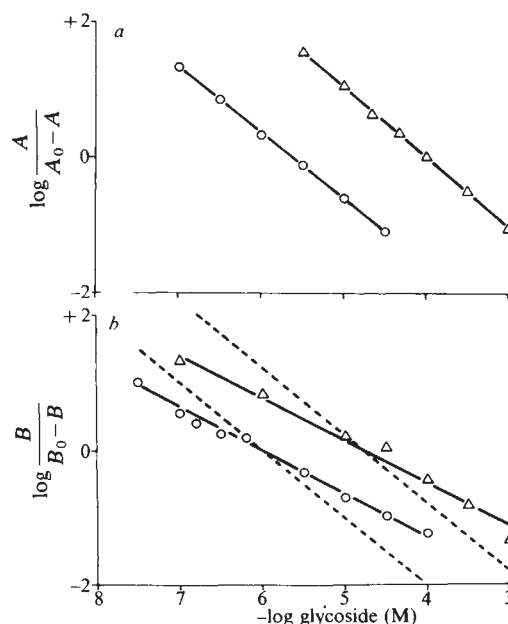


Fig. 1 Interaction of ouabain and dihydroouabain with guinea pig heart microsomes. The methods for microsomal $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ preparation and assay²⁰, and for ^3H -ouabain assay¹³ have been described elsewhere. *a*, $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ inhibition by cardiac glycosides. About 0.01 mg protein was incubated for 120 min in 2 ml medium containing 100 mM NaCl, 10 mM KCl, 3 mM MgCl_2 , 3 mM ATP, 1 mM EGTA, 20 mM Tris-maleate (pH 7.4, 37 °C) and various concentrations of glycoside. The released inorganic phosphate was measured by colorimetry²⁰. ATP hydrolysis was $<15\%$. Ordinate: $\log (A/A_0 - A)$ where A_0 is the control $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ activity and A the activity in the presence of ouabain (○) or dihydroouabain (△). Abscissa: $-\log \text{glycoside concentration}$. *b*, Displacement by ouabain or dihydroouabain of the specific ^3H -ouabain binding. About 0.24 mg protein was incubated in the conditions described above. The incubation medium was supplemented with 10^{-8} M ^3H -ouabain (19 Ci mmol^{-1} ; Amersham), and with 0.01 mM NaVO_3 , 5 mM inorganic phosphate, 3 mM creatine phosphate and 0.01 mg creatine kinase (25 U per mg) to limit ATP hydrolysis and keep the conditions constant⁸. The reaction was stopped by filtration on Whatman GF/F glass fibre filters. The filters were washed with 10 ml of chilled 0.25 M sucrose solution. The radioactivity trapped by the filters was measured as described previously¹³. The nonspecific binding was estimated in the presence of 1 mM ouabain and subtracted from the total. Ordinate: $\log (B/(B_0 - B))$ where B_0 is the control binding and B the binding in the presence of ouabain (○) or dihydroouabain (△). Abscissa: $-\log \text{glycoside concentration}$. The dashed lines represent theoretical curves with slope equal to 1. The points are means from three experiments. The departure from parallelism between experimental and theoretical curves was checked by a comparison of difference in slopes between the two curves (for ouabain, $t = 5.2$, $P < 0.001$; for dihydroouabain, $t = 8.30$, $P < 0.001$).

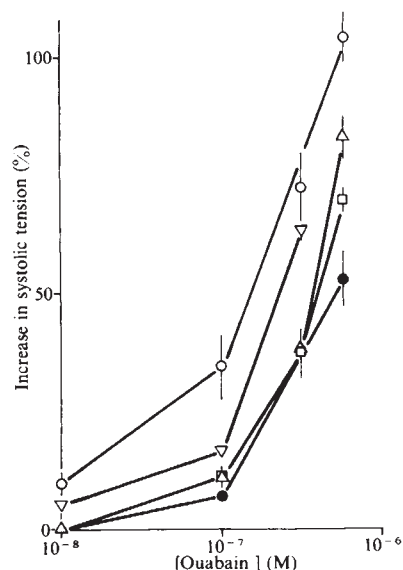


Fig. 2 Inotropic effect of ouabain in the presence or the absence of dihydroouabain. Adult albino guinea pigs (Dunkin Hartley strain, Hacking and Churchill Ltd) weighing 300–400 g were stunned and bled, their chests were opened and their hearts quickly excised. Left atria were dissected and suspended in an organ bath between platinum electrodes, under a preload tension of 500 mg. The physiological Tyrode solution contained (mM): NaCl, 137; KCl, 6; CaCl₂, 1.82; MgCl₂, 1.0; NaH₂PO₄, 0.417; NaHCO₃, 11.9; glucose, 5.5. It was equilibrated at 30°C with 95% O₂/5% CO₂. The atria were stimulated with rectangular 10-ms pulses (strength at least twice threshold) at a rate of 3.3 Hz. Recordings of the contractile activity were made by isometric lever using two strain gauges as part of a balanced bridge, the output of which was fed into a potentiometric recorder. After 30 min incubation in the absence (○) or the presence of dihydroouabain at concentrations of 10⁻⁸ M (Δ), 10⁻⁷ M (□), 3 × 10⁻⁷ M (□) or 10⁻⁶ M (●), ouabain was added to the medium. The inotropic effect was recorded at its maximum for each concentration of ouabain, that is, between 20 and 40 min after glycoside addition. The contractility of controls was not modified by dihydroouabain, except for a 10% increase with 10⁻⁶ M. The limits of s.e.m. are shown when they exceed the diameter of the symbol (*n* = 3–5). Ordinate: increase in systolic tension (%). Abscissa: ouabain concentration. Statistical analysis: difference in systolic contraction between randomized groups was checked by variance analysis before and after drug treatment. No statistical significance could be found before treatment. The effect of ouabain was analysed using a paired *t*-test. In the absence of dihydroouabain, ouabain at 10⁻⁸ M (*n* = 5) evoked a small but significant inotropic effect (0.02 < *P* < 0.05); at higher concentrations, ouabain effect was highly significant (*P* < 0.001). The influence of dihydroouabain pretreatment has been estimated by comparing the mean increase in contractility (expressed in g or in per cent of initial values) between groups treated by ouabain with or without different concentrations of the dihydro derivative. The *t*-test analysis showed the following results: ouabain 10⁻⁸ M (*n* = 5) + dihydroouabain 10⁻⁸ M (*n* = 3) not significant (NS), 10⁻⁷ M (*n* = 4) the response to ouabain was abolished; ouabain 10⁻⁷ M (*n* = 11) + dihydroouabain 10⁻⁸ M (*n* = 3) NS, 10⁻⁷ M (*n* = 6), 0.01 < *P* < 0.02, 10⁻⁶ M (*n* = 5), 0.001 < *P* < 0.01; ouabain 3 × 10⁻⁷ M (*n* = 10) + dihydroouabain 10⁻⁸ M (*n* = 3) NS, 10⁻⁷ M (*n* = 6), 0.001 < *P* < 0.01, 10⁻⁶ M (*n* = 5), 0.001 < *P* < 0.01; ouabain 6 × 10⁻⁷ M (*n* = 6) + dihydroouabain 10⁻⁷ M (*n* = 5), 0.02 < *P* < 0.05, 3 × 10⁻⁷ M (*n* = 4), 0.01 < *P* < 0.02, 10⁻⁶ M (*n* = 4), 0.02 < *P* < 0.05.

The experiments were performed on whole left atria stimulated by field electrodes in a medium containing 1.8 mM CaCl₂ so as to observe the action of low ouabain concentrations¹⁶. We compared the increase in atrial contractile force evoked by ouabain added after a preincubation with dihydroouabain concentrations varying from 10⁻⁷ M to 10⁻⁶ M, with that observed without this pretreatment. Figure 2 illustrates the relationship between ouabain concentration and the inotropic response of atria measured at its maximum and expressed as a percentage of controls, before glycoside. It shows how the ouabain inotropic effect was depressed by increasing the dihydroouabain concentration.

We have examined the action of ouabain on the Na⁺, K⁺-pump activity by measuring ⁴²K uptake in atria treated by ouabain in the presence or absence of dihydroouabain.

Figure 3 shows the relationship between the inhibition of ⁴²K uptake and the positive inotropic effect evoked by ouabain,

dihydroouabain or the combination of the two drugs. It confirms our previous report⁶ that for a given inhibition of active ⁴²K uptake, the inotropic effect of ouabain is higher than that evoked by dihydroouabain. The latter is directly related to Na⁺, K⁺-pump inhibition^{6,14}. For ouabain at 10⁻⁸, 10⁻⁷ and 3 × 10⁻⁷ M in the presence of 10⁻⁷ M dihydroouabain (which was without effect on ouabain inhibition of ⁴²K uptake), the relationship between the inhibition of ⁴²K uptake and the inotropic effect was similar to that observed with dihydroouabain at doses equiactive on ⁴²K uptake. This indicates that dihydroouabain did not antagonize the positive inotropic effect resulting from Na⁺, K⁺-pump inhibition, but only antagonized the additional mechanism believed to be specific to ouabain action. Figure 3 further illustrates how the antagonism exerted by dihydroouabain was surmounted by increasing ouabain concentrations.

It has been reported¹⁷ that (Na⁺ + K⁺)ATPase forms a complex with dihydroouabain that is less stable than the complex formed with ouabain. The higher stability of the ouabain-enzyme complex has been attributed¹⁸ to a two-point attachment of the lactone ring to the receptor: (1) through hydrogen bonding and (2) through an ionic bond between a partial positive charge at C-20 and an anionic group of the receptor. Ionic bonding does not occur with dihydroouabain because its lactone ring does not contain conjugated double bonds. This feature seems to be required for the intrinsic activity of ouabain at high-affinity sites that have been proposed to be responsible for the stimulation of the Na⁺, K⁺-pump.

In conclusion, the present results show that dihydroouabain reduces the inotropic action of ouabain but does not modify its inhibitory action on the Na⁺, K⁺-pump. As the inhibitory action of dihydroouabain on ouabain inotropy was overcome by increasing concentrations of this glycoside and as this action was observed at dihydroouabain concentrations devoid of direct pharmacological action, dihydroouabain can be considered as an antagonist of ouabain. The present results are consistent

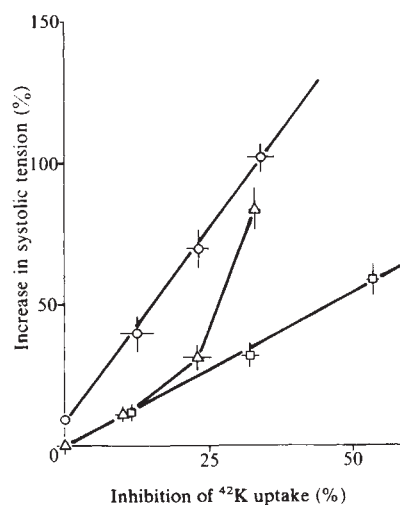


Fig. 3 Relationship between the inhibition of ⁴²K uptake and the increase in systolic tension of guinea pig atria. After incubation in the presence or absence of the glycosides as described in Fig. 2 legend, ⁴²K (Amersham) was added to the solution (5–10 nCi ml⁻¹); 10 min later, the atria were dipped for 5 s in a nonradioactive solution, blotted onto filter paper and pressed three times with a roller weighing 350 g. After weighing, each atrium was transferred to 0.2 ml of a solution composed of equal volumes of perchloric acid (37%, w/v) and H₂O₂. The tissue was dissolved by heating at 75°C for 30 min. After cooling, 5 ml of Aqualuma (Lumac) was added and the radioactivity of the samples was measured in a liquid scintillation counter. Inhibition of the ouabain-sensitive ⁴²K uptake, estimated as previously reported⁶, was evoked by: ○, ouabain at 10⁻⁸, 10⁻⁷, 3 × 10⁻⁷ and 6 × 10⁻⁷ M (*n* = 6); □, dihydroouabain at 10⁻⁶, 3 × 10⁻⁶ and 10⁻⁵ M (*n* = 8); Δ, a mixture of 10⁻⁷ M dihydroouabain preincubated 30 min before ouabain addition and the ouabain concentrations as stated above (*n* = 8). The limits of s.e.m. are shown when they exceed the diameter of the symbol. Ordinate: increase of systolic tension as % of the control. Abscissa: % inhibition of ouabain-sensitive ⁴²K uptake.

with previous reports that ouabain inotropic action cannot be entirely^{5-7,19} accounted for by the glycoside-evoked changes in Na^+ and K^+ gradients.

In preliminary experiments, these observations have been extended to other preparations (sometimes dihydroouabain 3×10^{-6} M being required). Such experiments could contribute

Received 15 February; accepted 2 September 1982.

1. Bonting, S. L. in *Membrane and Ion Transport* (ed. Bittar, E. E.) 257-363 (Wiley-Interscience, London, 1970).
2. Schwartz, A., Lindenmayer, G. E. & Allen, J. C. *Pharmac. Rev.* **27**, 1-134 (1975).
3. Godfraind, T. & Ghyssels-Burton, J. *Nature* **265**, 165-166 (1977).
4. Langer, G. A. *Biochem. Pharmac.* **30**, 3261-3264 (1981).
5. Browning, D. J., Guarneri, T. & Strauss, H. C. *J. clin. Invest.* **68**, 942-956 (1981).
6. Godfraind, T. & Ghyssels-Burton, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 3067-3069 (1980).
7. Ghyssels-Burton, J. & Godfraind, T. *Br. J. Pharmac.* **66**, 175-184 (1979).
8. Godfraind, T., De Pover, A. & Tona Lutete, D. *Biochem. Pharmac.* **29**, 1195-1199 (1980).
9. De Pover, A., Godfraind, T. & Tona Lutete, D. *Br. J. Pharmac.* **72**, P513-P514 (1981).

to a better understanding of the role of the inhibition of the Na^+ , K^+ -pump in the pharmacological action of ouabain.

We thank Dr M. Wibo for reading the manuscript, Mr A. Goffin for technical assistance and Mr P. Simon for the drawings. This work was supported by grant 3'9001.79 of Fonds de la Recherche Scientifique Médicale (Belgium).

10. Wellsmith, N. V. & Lindenmayer, G. E. *Circulation Res.* **47**, 710-720 (1980).
11. Erdmann, E., Philipp, G. & Scholz, H. *Biochem. Pharmac.* **29**, 3219-3229 (1980).
12. Schwartz, A. et al. *Proc. 3rd int. Conf. Na, K-ATPase* (Academic, New York, in the press).
13. De Pover, A. & Godfraind, T. *Biochem. Pharmac.* **28**, 3051-3056 (1979).
14. Lee, C. O., Kang, D. H., Sokol, J. H. & Lee, K. S. *Biophys. J.* **29**, 315-330 (1980).
15. Pocock, G. J. *Physiol., Lond.* **307**, 38P (1980).
16. Grupp, G., Grupp, I. L., Ghyssels-Burton, J., Godfraind, T. & Schwartz, A. *J. Pharmac. exp. Ther.* **220**, 145-151 (1982).
17. Clark, A. F., Swanson, P. D. & Stahl, W. L. *J. biol. Chem.* **250**, 9355-9359 (1975).
18. Thomas, R., Boutagy, J. & Gelbart, A. *J. Pharmac. exp. Ther.* **191**, 219-231 (1974).
19. Noble, D. *Cardiovasc. Res.* **14**, 495-514 (1980).
20. Godfraind, T., De Pover, A. & Verbeke, N. *Biochim. biophys. Acta* **481**, 202-211 (1977).

Hydrogen ion currents and intracellular pH in depolarized voltage-clamped snail neurones

R. C. Thomas & R. W. Meech

Department of Physiology, University of Bristol, University Walk, Bristol BS8 1TD, UK

Until now the only reported effect of depolarization on the intracellular pH (pH_i) of excitable cells is an acidification of the cell cytoplasm^{1,2}. It seems unlikely that this could be a direct effect of membrane potential because pH_i is known to be regulated by an electroneutral mechanism^{3,4} and in most cells H^+ ions are not in equilibrium with the membrane potential (E_m). In any case the membrane conductance to H^+ ions would be expected to be small because they are at such low concentrations on either side of the cell membrane. But it is possible that the H^+ ion permeability of the membrane increases on depolarization just like that of other ions in the bathing medium (Na^+ , K^+ and Ca^{2+} for example). To test this idea we have made pH_i measurements on molluscan neurones under voltage-clamp. Our findings, presented here, provide evidence for a large increase in H^+ ion permeability in depolarized cells. We suggest that this increase in proton conductance may be the basis for the 'nonspecific' currents previously described in perfused molluscan neurones^{5,6} and we assess the physiological significance of this newly discovered pathway.

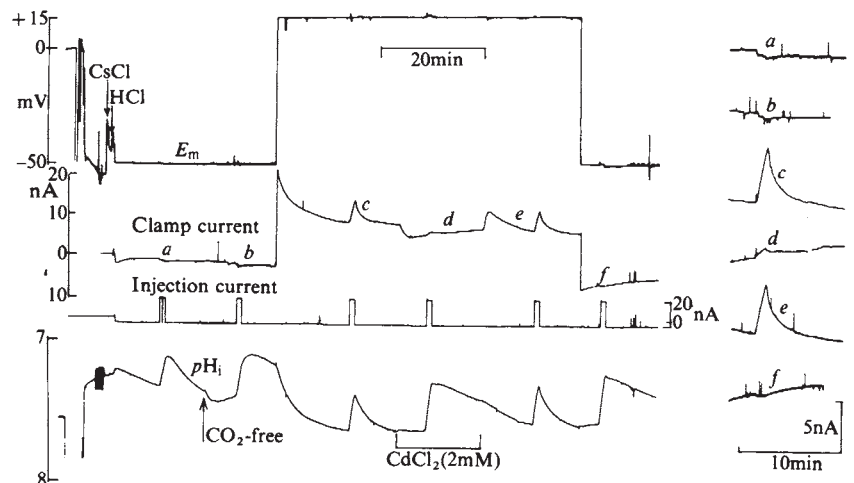
Experiments were done on exposed cells in isolated suboesophageal ganglia from the snail, *Helix aspersa*. The neuronal cell body was penetrated with four microelectrodes. The first was to measure pH_i . The second was a reference liquid ion-exchanger electrode to measure E_m ⁸. The third, a micropipette

filled with 3 M KCl, CsCl or $(\text{Cs})_2\text{SO}_4$, supplied feedback current to voltage-clamp the cell⁹. The fourth was for iontophoretic HCl injection. The preparation was superfused with saline buffered to pH 7.5 either with HEPES alone or with HEPES and 4 mM NaHCO_3 . The bicarbonate saline was equilibrated with 0.5% CO_2 in air.

Figure 1 shows an experiment in which HCl was injected at two different holding potentials, -50 mV and +15 mV. After the first injection, at -50 mV, pH_i recovered within 10 min. This pH_i recovery involves a mechanism which exchanges external Na^+ and HCO_3^- ions for internal Cl^- and H^+ ions³. It was inhibited by changing the bathing medium to a bicarbonate-free saline. The next HCl injection produced a slightly larger decrease in pH_i because the buffering capacity of the cytoplasm is smaller in CO_2 -free saline^{10,11}. Recovery from the injection was slow and before it was complete we depolarized the cell to +15 mV. This caused a rapid unexpected increase in pH_i to a steady level of 7.6. (The normal pH_i is between 7.3 and 7.5.) After a third HCl injection the pH_i rapidly returned to this new steady level. Despite the lack of bicarbonate the recovery was faster than the control in normal saline. In other experiments normal pH_i regulation was abolished by using Na-free saline or the inhibitor SITS (4-acetamide-4'-isothiocyanatostilbene-2,2'-disulphonic acid) and the cell was depolarized either under voltage-clamp or with isotonic KCl¹². In each case there was a rapid recovery of pH_i following an acid load.

At depolarized potentials pH_i recovers from acid injections by a mechanism clearly different from that operating normally. Further evidence for this is that in depolarized cells pH_i recovery is sensitive to heavy metals in the bathing medium. Figure 1 shows that the recovery of pH_i following an acid load at +15 mV was slowed by 2 mM CdCl_2 . When the Cd^{2+} ions were removed there was a significant increase in the rate of pH_i recovery. Preliminary tests with other heavy metals showed that 1 mM CuCl_2 or 1 mM LaCl_3 had a similar effect.

Fig. 1 The effect of HCl injections on the intracellular pH (pH_i) and clamp current of a voltage-clamped snail neurone. The four traces are pen-recordings of membrane potential (E_m), clamp current, injection current and pH_i . Enlargements of the clamp current during the six acid injections are shown at the right. The general experimental procedures have been described previously^{3,7,11}. The clamp time constant was such that the cell could fire action potentials at the resting potential⁹ while at positive potentials there was good voltage control. But the cells survived equally well with a rapidly responding clamp, that is, when there were no action potentials at any potential. In most experiments the voltage-clamp current electrode was filled with CsCl or Cs_2SO_4 so that Cs^+ ions were injected when the cell was depolarized. Intracellular Cs reduces the outward K^+ current and hence gives improved voltage control. The arrows above the E_m record indicate the point at which the current microelectrodes were inserted.



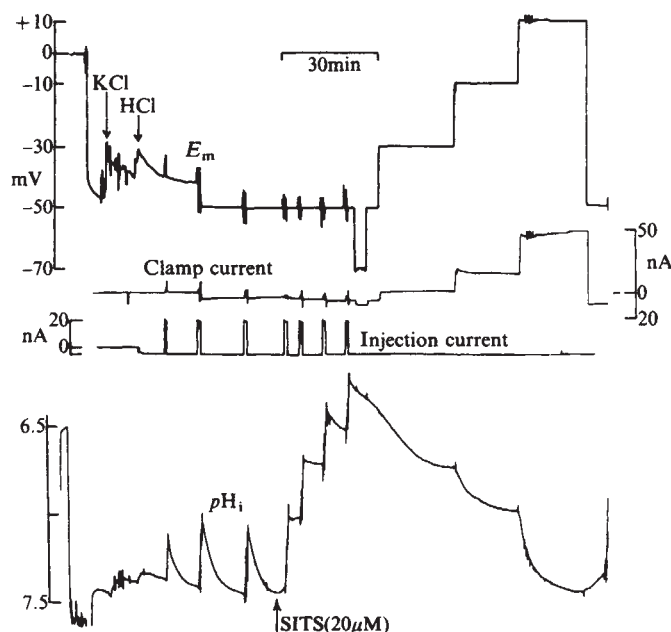


Fig. 2 The intracellular pH (pH_i) at different membrane potentials (E_m) in a voltage-clamped cell. After the first three injections the normal pH_i regulating system was blocked by SITS. The voltage-clamp current electrode was filled with KCl.

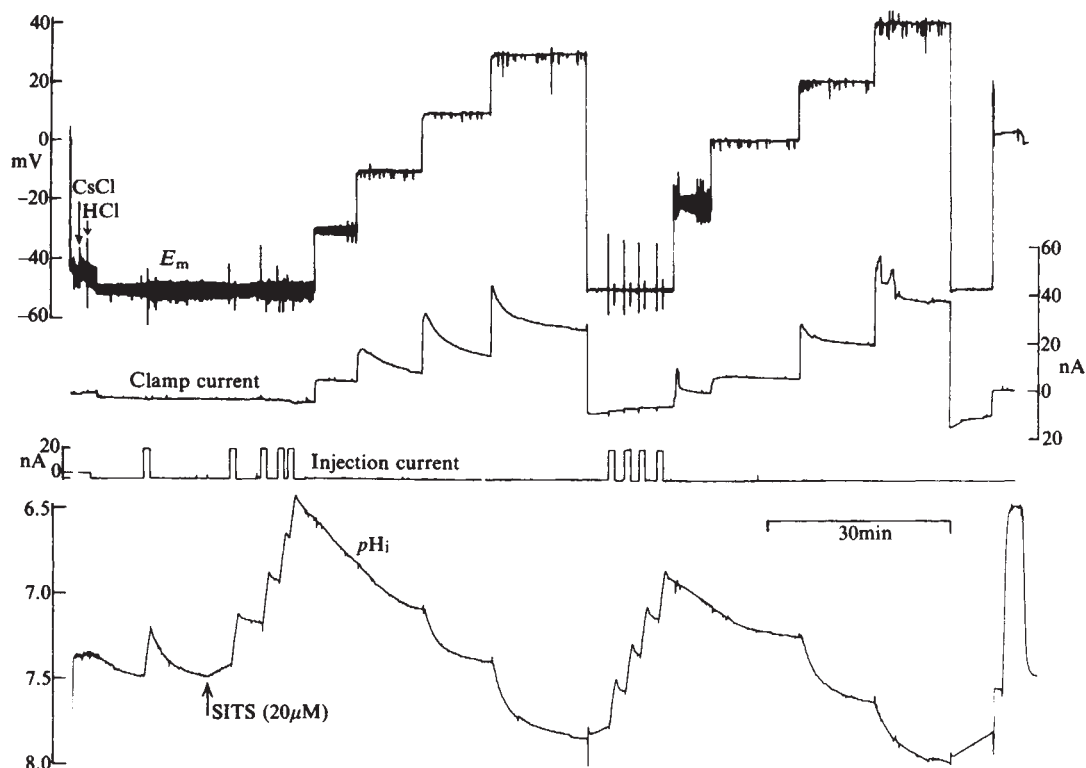
We tested the way in which the membrane potential level influenced the pH_i recovery in the experiment shown in Fig. 2. In this cell pH_i regulation was first abolished by bathing the cell in saline containing $20 \mu M$ SITS. Then four separate iontophoretic injections of HCl took pH_i to 6.23. Note that as pH_i was made more acid its rate of recovery increased, but never approached that of the control. The pH_i recovery was unaffected by hyperpolarizing the membrane to -70 mV for 5 min, but its rate increased when the membrane was depolarized to -30 mV. After about 36 min the pH_i reached a steady level at pH 6.74 but a further 20 mV depolarization produced a further increase in pH_i until it reached another steady level at 6.98. At $+10$ mV

pH_i became steady at pH 7.44. Heavy metals inhibit these potential-dependent pH_i shifts (not illustrated). In preliminary tests we find that they are blocked by $CdCl_2$, $LaCl_3$, $ZnCl_2$ (1 mM) or $CoCl_2$ (10 mM) and that they are slower in 4 mM Ruthenium red.

In experiments of this kind using KCl-filled micropipettes we were unable to test the effect of depolarizations beyond about $+20$ mV because the cells did not survive for long enough. When we used CsCl the outward currents were greatly reduced and, perhaps as a consequence, the cells survived for longer periods. In the experiment shown in Fig. 3 the membrane potential was increased to $+37.5$ mV in 20-mV steps. With each step beyond -10 mV, there was a rapid pH_i increase to a steady level such that the H^+ ion equilibrium potential also changed by about 20 mV. The relationship between E_m and the steady pH_i level is shown in Fig. 4. The solid line drawn through the points shows the slope expected if pH_i changes by 1 pH unit for a 58-mV change in E_m . The interrupted line shows the expected result if the membrane potential was at the hydrogen ion equilibrium potential, E_H . If this were the case pH_i would equal the external pH (that is, pH 7.5) when the membrane potential was held at 0 mV. In fact, as Fig. 4 shows, pH_i is close to 7.3 at 0 mV. In four experiments the average difference between the internal and external pH at 0 mV was 0.33 ± 0.09 ($\pm$ s.d.) pH unit. In the absence of SITS the difference was 0.07 ± 0.04 pH unit ($n = 4$). It is possible that this is because of unstirred layers near the cell membrane.

A satisfactory explanation for the potential dependence of pH_i seen at depolarized potentials is that the membrane becomes so highly permeable to H^+ (and/or OH^-) ions that their concentration inside the cell is determined simply by the external pH and the potential across the cell membrane. H^+ ions leave the cell cytoplasm until their gradient across the cell membrane is in equilibrium with E_m . It is likely that most of the H^+ ions leave the cytoplasm by crossing the cell membrane and if so, it may be possible to detect a large outward current. Indeed, such a current can be seen in the clamp current trace in Fig. 1. Injections *a*, *b* and *f* which were carried out at -50 mV produced no obvious changes in the current trace but *c* and *e* produced a large outward current which declined in an exponential fashion once the injection was terminated. The presence of 2 mM Cd^{2+} , which blocked rapid pH_i recovery, also blocked

Fig. 3 Relationship between membrane potential (E_m) and intracellular pH (pH_i) in a neurone injected with Cs. After the first HCl injection the normal pH_i regulating system was blocked by SITS. At the end of the experiment the pH electrode response to calibration with pH 6.5 saline is shown. The apparent noise on the E_m record is caused by action potentials, which are greatly reduced by the long time constant of the pen-recorder and are prolonged by Cs^+ injection under voltage-clamp.



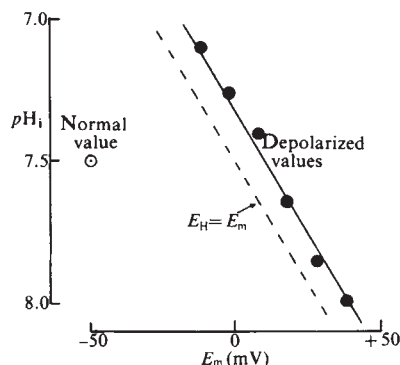


Fig. 4 Graph showing the relationship between membrane potential (E_m) and steady-state intracellular pH (pH_i) from the experiment of Fig. 3. ●, Steady pH_i values taken after the cell had been treated with SITS and depolarized. Solid line drawn by eye with slope of 58 mV per pH unit to give best fit to points. Dashed line shows relationship expected of pH_i determined by purely passive distribution of H^+ ions across cell membrane with no surface charges or net flux of H^+ , OH^- or HCO_3^- ions.

the appearance of the outward current (see injection *d*). The time constant of the decline in the injection dependent current is about 1.4 min, which is half the time constant of the associated pH_i recovery. One possible reason for this difference is that our measurements of pH_i changes are limited by the response time of the pH electrode, and its location in the cell.

If all the injected H^+ (and OH^-) ions leave the cell by crossing the cell membrane and there are no other ion movements, then the injection-induced charge movement should equal the charge used to inject the H^+ ions in the first place. The charge movement generated by injection *c* was 5.97×10^{-7} C which was about 66% of the charge used to inject HCl (assuming a transport number for H^+ of 0.94¹⁰). It is possible that the H^+ movement is coupled to the movement of other ions or that up to a third of the injected H^+ ions are taken into intracellular organelles¹³. Potential-dependent shifts in pH_i could be produced even after 80 min in saline containing *n*-methylglucamine methane sulphonate and HEPES buffer only.

Does this newly discovered H^+ ion permeable pathway operate during the normal activity of the cell? To be of any physiological significance the ions would have to move during the short time that the membrane potential is more positive than about -20 mV during an action potential. If they do it should be possible to identify an outward H^+ current during brief depolarizing pulses. Just like the HCl injection-induced current (Fig. 1), it should be blocked by many of the agents which block calcium channels (Cd^{2+} , Co^{2+} , La^{3+} etc., see ref. 14 for review), it would be expected to reverse near 0 mV and should be sensitive to changes in external pH. Byerly and Hagiwara⁶ have recently described an apparently 'nonspecific' outward current in perfused *Limnaea* neurones which has all these characteristics. A similar outward current has been described in perfused *Helix* neurones⁵. In this case its current-voltage characteristic curve shifts by +15 mV for a unit change in external pH¹⁵. Both currents are rapidly activated. If the nonspecific current is an outward H^+ current as we suggest it could have a significant effect on pH values close to the membrane in actively firing cells and may partly compensate for the rapid acidification described by Ahmed and Connor¹.

The H^+ currents we describe could perhaps be nonspecific in the sense that they flow through several different types of activated channel. For this reason it seems worthwhile to test for pH sensitivity the nonspecific calcium-dependent single channel currents which have been recently described in heart muscle cells¹⁶. H^+ channels may be more widespread than hitherto suspected.

This work was supported by the MRC and by the Wellcome Trust. Preliminary experiments by R.W.M. were carried out in the Physiology Department, University of Utah, supported by

NIH grant EY 02290 and by R.C.T. in the Physiology Department of Yale University partly supported by USPHS grant AM 17322. We thank Michael Rickard and Toni Gillett for technical help, and Bruce Matthews and Sandy Harper for comments on the manuscript.

Received 19 July; accepted 1 September 1982.

1. Ahmed, Z. & Connor, J. A. *J. gen. Physiol.* **75**, 403-426 (1980).
2. Abercrombie, R. F. & Roos, A. *J. Physiol., Lond.* **319**, 65P-66P (1981).
3. Thomas, R. C. *J. Physiol., Lond.* **273**, 317-338 (1977).
4. Roos, A. & Boron, W. F. *Physiol. Rev.* **61**, 296-434 (1981).
5. Kostyuk, P. G. & Krishtal, O. A. *J. Physiol., Lond.* **270**, 545-568 (1977).
6. Byerly, L. & Hagiwara, S. *J. Physiol., Lond.* **322**, 503-528 (1982).
7. Thomas, R. C. *Ion-Sensitive Intracellular Microelectrodes* (Academic, London, 1978).
8. Thomas, R. C. & Cohen, C. J. *Pflügers Arch. ges. Physiol.* **390**, 96-98 (1981).
9. Thomas, R. C. *J. Physiol., Lond.* **201**, 495-514 (1969).
10. Thomas, R. C. *J. Physiol., Lond.* **255**, 715-735 (1976).
11. Meech, R. W. & Thomas, R. C. *J. Physiol., Lond.* **298**, 111-129 (1980).
12. Thomas, R. C. *J. Physiol., Lond.* **296**, 77P (1979).
13. Mitchell, P. & Moyle, J. *Biochem. J.* **104**, 588-600 (1967).
14. Hagiwara, S. & Byerly, L. A. *Rev. Neurosci.* **4**, 69-125 (1981).
15. Doroshenko, P. A., Kostyuk, P. G. & Tsyndrenko, A. Ya. *Neirofiziolgiya* **10**, 645-653 (1978). (Translated in *Neurophysiology* **10**, 470-477; 1978).
16. Colquhoun, D., Neher, E., Reuter, H. & Stevens, C. F. *Nature* **294**, 752-754 (1981).

Intracellular Cl^- accumulation reduces Cl^- conductance in inhibitory synaptic channels

Michael R. Gold & A. R. Martin

Department of Physiology, School of Medicine, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

Glycine increases Cl^- conductance in neurones of the lamprey central nervous system and mimics the natural inhibitory transmitter pharmacologically^{1,2}. We have used current noise produced by glycine to examine in more detail the characteristics of inhibitory channels in Müller cells in the brain stem reticular formation. The channels have large conductances and mean open times consistent with large amplitudes and long durations of spontaneously occurring inhibitory synaptic currents³. We now show that, unlike any post-junctional channels reported so far, their conductance decreases rapidly with increasing intracellular concentration of the permeant ion. This surprising behaviour is inconsistent with constant field theory^{4,5} and also with a single-site pore model such as proposed for cationic channels at the motor endplate⁶, both of which predict an increase in conductance with concentration. In addition, the decrease in conductance cannot be explained quantitatively by a two-site, single-file pore model such as proposed for K^+ channels⁷ and Na^+ channels⁸ in nerve and for gramicidin channels⁹. This property of the inhibitory channel may be functionally important in preventing intracellular Cl^- accumulation during periods of intense synaptic activity when inhibition might otherwise convert progressively to excitation as the Cl^- equilibrium potential became more and more positive.

Brains were isolated from adult lampreys (*Petromyzon marinus* or *Lampetra lamottenii*) anaesthetized with tricaine methylsulphonate (0.5 g l⁻¹), then superfused with an oxygenated saline solution containing (mM): 104.5 NaCl, 2.0 KCl, 5.0 CaCl₂, 5.0 MgCl₂, 4.0 glucose, 2.0 HEPES, and adjusted to pH 7.4. Tetrodotoxin (1.0 μM) and 4-aminopyridine (1.0 mM) were added to prevent spontaneous activity and to reduce delayed rectification. The bathing solution was cooled to ~4 °C. Müller cells were impaled with two electrodes for voltage-clamping. Both the voltage and current electrodes were filled with either 3 M K^+ -acetate or 3 M KCl and had resistances in the range 20-30 MΩ. The voltage-clamp amplifier had a bandpass of 0-1.8 kHz and a maximum gain of 10,000. Glycine

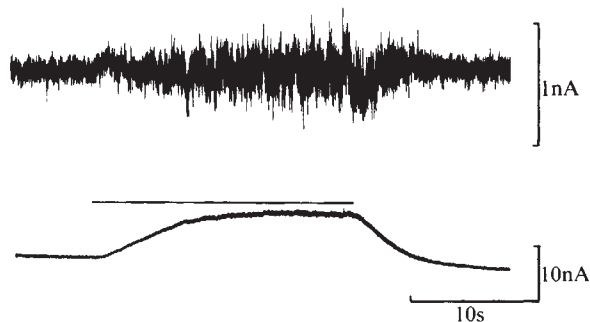


Fig. 1 Example of current noise produced by a 20-s pulse of glycine to a Müller cell. Agonist application (bar) produced a peak outward current of 6.9 nA (lower trace). The accompanying increase in baseline noise is shown in the upper trace at higher gain through an a.c. amplifier with a bandpass of 0.75–140 Hz (half-power points).

was applied to the cells near the base of the dendrites by pressure pulses to a third pipette filled with a 0.2–1.0 M solution of the agonist; the resulting membrane currents were amplified by a low gain d.c. and higher gain a.c. amplifier and stored on two channels of an FM tape recorder for later analysis. Membrane potentials were clamped 10–20 mV away from the reversal potential for the glycine response in either the positive or negative direction.

Figure 1 shows a typical cell response to glycine. Application of the agonist for 20 s produced an outward current of ~10 nA (lower trace) and a marked increase in baseline noise (upper trace). The magnitudes of the frequency components of this excess noise, as represented by their spectral density distribution, were used to determine the conductance and mean open time of the glycine-activated channels¹⁰. Spectral distributions were calculated as the Fourier transforms of 1,024-point samples from the experimental records digitized at 512 or 1,024 points s⁻¹. Spectra of baseline fluctuations before glycine application were subtracted from those obtained during a steady-state response, and 8–32 such difference spectra were averaged to obtain a final distribution. All spectra were consistent with a single distribution of the form $S(f) = S(0)[1 + (f/f_c)^2]^{-1}$, suggesting that glycine opened channels whose lifetimes were distributed exponentially with a mean duration $\tau = 1/2\pi f_c$. The individual channel conductance (γ) was calculated as $\gamma = S(0)/4\tau I\Delta V$, where I is the mean current through the membrane, ΔV the difference between the reversal potential and the holding potential and $S(0)/4\tau$ the area under the distribution (that is, the total variance due to the agonist action).

Spectral distributions from two cells are shown in Fig. 2. The results in Fig. 2a were obtained with acetate-filled electrodes. The reversal potential for the response was -72 mV and the membrane was clamped at -87 mV. Channel conductance was calculated to be 94 pS and the mean channel open time 29 ms. In 14 experiments of this type, the overall mean channel conductance was 73 ± 15 pS (mean $\pm$ s.d.) and the mean channel open time 33 ± 9 ms. The average reversal potential was 66 ± 7 mV. Neither conductance nor open time was significantly dependent on membrane potential. When Cl⁻-filled electrodes were used, the mean channel conductance was markedly less than with acetate electrodes, apparently as a result of loading the cell with Cl⁻. An example of the spectral density distribution from one such cell is shown in Fig. 2b: the reversal potential for the response was -44 mV, reflecting the increase in intracellular Cl⁻, and the membrane was clamped at -54 mV. Mean channel conductance was calculated to be 22 pS and channel open time 35 ms. In 10 such experiments, reversal potentials ranged from -35 to -60 mV and channel conductances from 7 to 43 pS. In four additional experiments acetate electrodes were used and intracellular Cl⁻ was increased in a different way, that is, by increasing extracellular K⁺ concentration². This

again resulted in a positive shift in mean reversal potential (to -47 mV) and channel conductance was reduced to a mean of 30 ± 9 pS. In contrast to the marked changes in γ , τ was unaffected by intracellular Cl⁻ concentration.

The reversal potential for the glycine response was used to calculate internal Cl⁻ concentration, assuming that the response reversed at the Cl⁻ equilibrium potential. The resulting relationship between channel conductance and intracellular Cl⁻ concentration is plotted in Fig. 3. The filled circles represent the results obtained with acetate electrodes, the crosses with Cl⁻ electrodes and the open circles with acetate electrodes and increased external K⁺. Clearly, the channel conductance varies inversely with some function of the internal Cl⁻ concentration, $[Cl^-]_i$. If we arbitrarily assume an inverse power function, $\gamma = \alpha[Cl^-]_i^{-n}$, then $n = 1.5$ gives a reasonable fit to the results (solid line). This result is inconsistent with constant field theory^{4,5} and with a single-site pore model⁶, both of which predict an increase in conductance with concentration. Two-site, single-file pore models, however, do predict a decrease in conductance as ionic concentrations increase on either side of the membrane⁷⁻⁹. However, when such models are examined in detail it can be shown that the relationship between conductance and concentration at the high concentration limit assumes the form

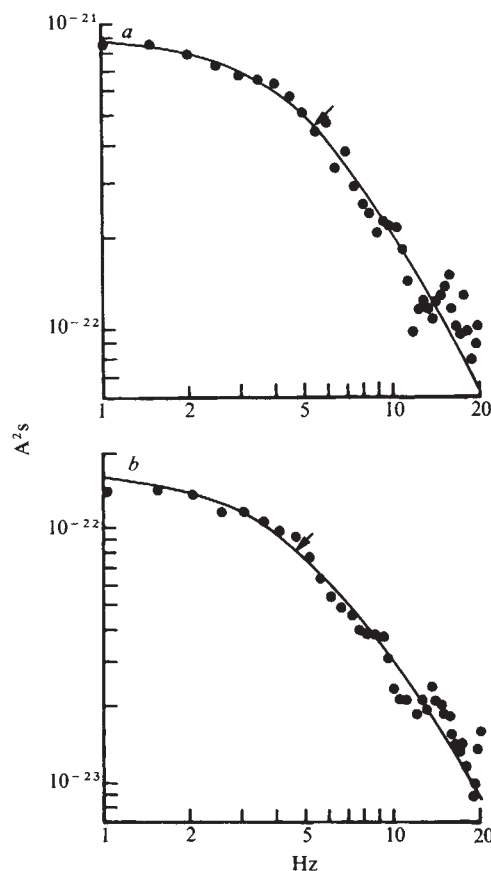


Fig. 2 Power density ($A^2 s$) as a function of frequency (Hz) for two different experimental conditions. Smooth curves are of the form $S(f) = S(0)[1 + (f/f_c)^2]^{-1}$ where $S(f)$ is the power density and f the frequency. These were fitted to the results by eye by selecting asymptotic power density, $S(0)$ and corner frequency, f_c . *a*, Experiment in normal solution with acetate electrodes. Points are the means of 16 difference spectra smoothed by weighted averaging of adjacent points in groups of five. Reversal potential -72 mV, holding potential -87 mV, mean current 5.5 nA. Curve fit with $S(0) = 9 \times 10^{-22} A^2 s$, $f_c = 5.5$ Hz. Estimated channel conductance $\gamma = 94$ pS, mean open time $\tau = 29$ ms. *b*, Average of 32 difference spectra from experiment with Cl⁻-filled electrodes (increased internal Cl⁻). Reversal potential -44 mV, holding potential -54 mV, mean current 4.8 nA. $S(0)$ selected as $1.5 \times 10^{-22} A^2 s$, f_c 4.5 Hz; γ and τ estimated as 22 pS and 35 ms, respectively.

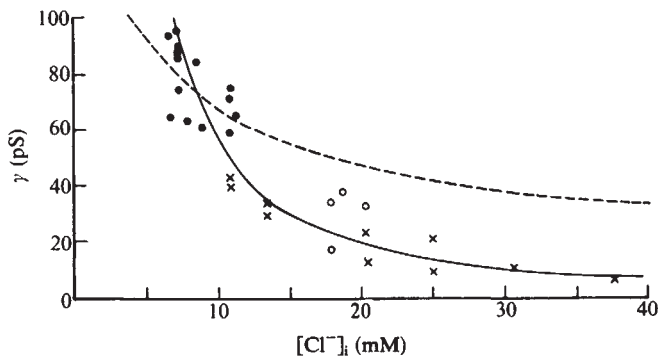


Fig. 3 Relationship between channel conductance (γ) and internal Cl^- concentration, $[\text{Cl}^-]_i$. $[\text{Cl}^-]_i$ was calculated from the reversal potential for the glycine response, E_r , according to the Nernst relation $[\text{Cl}^-]_i = [\text{Cl}^-]_o \exp(E_r F/RT)$, where $[\text{Cl}^-]_o$ is the external Cl^- concentration and R , T and F have their usual meanings. With normal solution and acetate electrodes (●), mean internal Cl^- was 8.4 mM and mean channel conductance 73 pS. The solid curve is drawn through this point according to the relation $\gamma = \alpha[\text{Cl}^-]_i^{-1.5}$; the broken line is the relationship predicted by a two-site, single-file model of the Cl^- channel (see text). ×, Points obtained with Cl^- -filled electrodes; ○, experiments with acetate electrodes and increased (10 mM) external K^+ .

$\gamma = \beta(C_i C_o)^{-0.5}$, where C_i and C_o are internal and external ionic concentrations, respectively, and β is a constant. Thus, with respect to changes in internal concentration, we would expect the relationship to be an inverse power function with $n = 0.5$. This type of relation (broken line in Fig. 3), is clearly inconsistent with the present results. Studies of additional models are under way. Note that an apparent reduction in channel conductance can be obtained due to receptor saturation by the agonist¹⁰. This was not the case in the present experiments as the estimated conductances were independent of steady-state currents (and hence of glycine concentration at the receptors) over the range of currents used.

The behaviour reported here for individual channel conductances has also been observed with spontaneously occurring inhibitory synaptic currents—these become smaller as internal Cl^- concentration increases³. Thus the reduction in conductance is related specifically to postsynaptic channels. The functional consequence of such behaviour is that any Cl^- accumulation within the cells will result in a decrease in inhibitory conductance and a consequent reduction in further Cl^- entry. Although inhibitory synaptic activity alone would be expected to lead to very little slow entry (the membrane simply being hyperpolarized to near the Cl^- equilibrium potential), combined excitatory and inhibitory activity could, in theory, lead to a considerable increase in internal Cl^- concentration. The negative feedback of concentration on conductance may be important in limiting such accumulation. Otherwise inhibitory inputs might, over a period of time, invert to become excitatory as the Cl^- equilibrium potential became more and more positive.

This work was supported by research grant NS-09660 and fellowship NS-06283 from the NIH.

Denervation of newborn rat muscles does not block the appearance of adult fast myosin heavy chain

Gillian S. Butler-Browne*, Lawrence B. Bugaisky*, Sylvia Cuénoud*, Ketty Schwartz† & Robert G. Whalen*

* Département de Biologie Moléculaire, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris, France

† INSERM Unit 127, Hôpital Lariboisière, 41 Bd. de la Chapelle, 75010 Paris, France

Several observations, both *in vivo* and *in vitro*, have indicated that the development and maturation of mammalian skeletal muscle fibres is influenced by nerve-muscle interactions. Morphological maturation of newly regenerated adult mouse muscle fibres in an organotypic nerve-muscle culture system depends on the presence of spinal cord neurones¹. Sciatic nerve transection in newborn rats has been shown to modify the development of the histochemical and contractile properties of the denervated muscles²⁻⁴. In addition, neural influences are important for the appearance of certain of the myosin small subunits^{5,6}. It has been proposed that the nerve also controls the changes in myosin heavy chain isozymes appearing during development^{5,7-9}. One such transition occurs in rat muscle where the neonatal form of myosin heavy chain is replaced by the adult form during the second post-natal week⁸. Here we demonstrate that innervation of the rat gastrocnemius muscle (a fast-contracting muscle in the adult) is not required for the appearance of the adult form of myosin heavy chain.

Muscles of the lower hind leg were denervated by removing a 4–5 mm piece of the sciatic nerve of rats 1 week after birth. The presence of the various myosin heavy chain isozymes was assessed by an immunocytochemical technique using antibodies specific for the various isozymes. In the gastrocnemius muscle of 7–8 day control (unoperated) rats, no reactivity is seen with the anti-fast serum whereas anti-neonatal immunoglobulins stain all fibres that react histochemically as type II (alkali-stable ATPase) (results not shown). From about 2 weeks after birth the number of fibres that stain with the anti-neonatal immunoglobulin begins to decrease, either in control muscles of unoperated rats or in the innervated muscles of operated rats. At 3 weeks of age, in a region of the innervated muscle composed mainly (>90%) of type II fibres, staining serial sections with antibodies against fast and neonatal myosin reveals that about 50% of the fibres react with either fast (~25%) or neonatal (~25%) antibodies whereas ~40% react with both (Fig. 1a, b). By 4–5 weeks after birth in the innervated muscle of operated rats, the percentage of fibres staining with anti-neonatal immunoglobulin alone decreases considerably to <10% and the percentage staining with anti-fast serum alone increases to ~80%.

In the denervated muscle of 22-day old animals (2 weeks after denervation), there is variation in fibre size, with many atrophic fibres. Most fibres irrespective of their diameter react with both fast and neonatal antibodies (Fig. 1c, d), in contrast to the innervated muscles (Fig. 1a, b). By 35 days of age, the staining for neonatal myosin is considerably decreased in the denervated muscle while staining with the anti-fast myosin serum is still present (Fig. 1e, f). A population of small fibres is also seen (Fig. 1f) which reacts strongly with the anti-neonatal immunoglobulin; these small fibres are not seen in control muscles nor in the innervated muscles of operated rats (not shown).

As shown by others¹⁰, many fibres in the fetal (18–20 days gestation) and newborn rat EDL muscle (a future fast-contracting muscle) react with an antibody to slow myosin. In these studies the possible cross-reactivity of the anti-slow immunoglobulin with embryonic or neonatal myosin was not

Received 29 June; accepted 6 September 1982.

1. Homma, S. & Rovainen, C. M. *J. Physiol., Lond.* **279**, 231–252 (1978).
2. Matthews, G. & Wickelgren, W. O. *J. Physiol., Lond.* **293**, 393–415 (1979).
3. Gold, M. R. & Martin, A. R. *Proc. physiol. Soc.*, 16–17 July (1982).
4. Goldman, D. E. *J. gen. Physiol.* **27**, 37–60 (1943).
5. Hodgkin, A. L. & Katz, B. *J. Physiol., Lond.* **108**, 37–77 (1949).
6. Lewis, C. A. *J. Physiol., Lond.* **286**, 417–445 (1979).
7. Hille, B. & Schwarz, W. *J. gen. Physiol.* **72**, 409–422 (1978).
8. Begenisich, T. B. & Cahalan, M. D. *J. Physiol., Lond.* **307**, 217–242 (1980).
9. Urban, B. W., Hladky, S. B. & Haydon, D. A. *Biochim. biophys. Acta* **602**, 331–354 (1980).
10. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655–691 (1973).

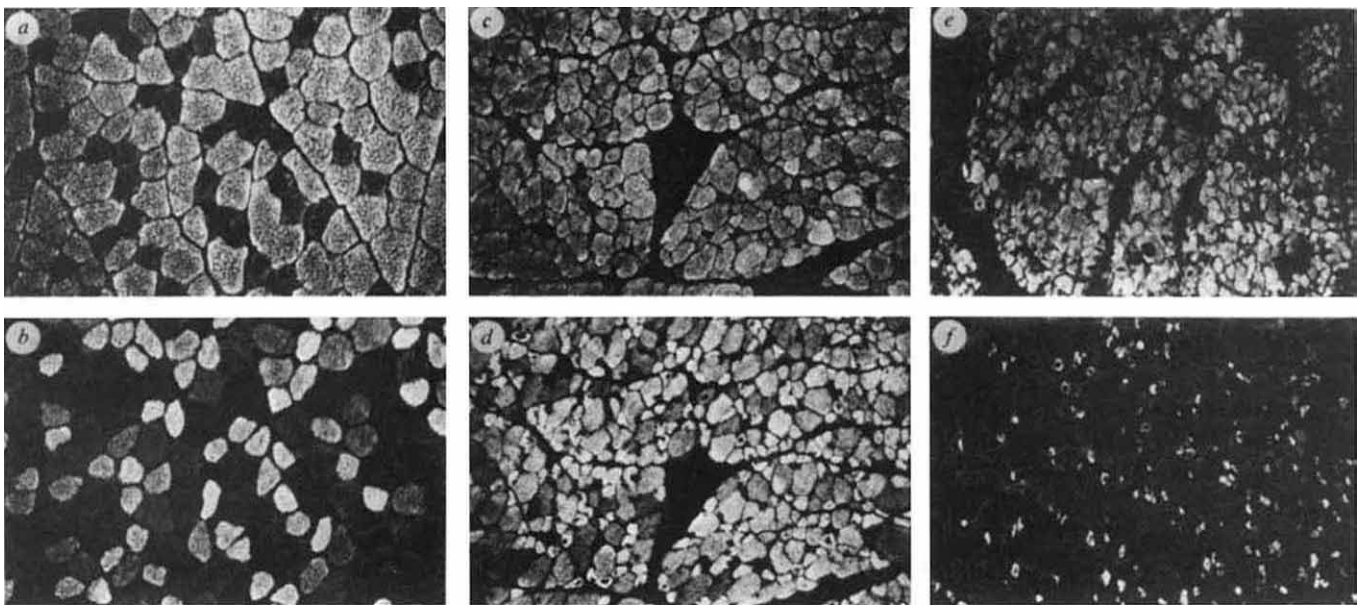


Fig. 1 Indirect immunofluorescence using antibodies to adult fast and neonatal myosin on cryostat sections of denervated and contralateral innervated rat gastrocnemius muscles. The antiserum to adult fast myosin was prepared by injecting guinea pigs with SDS-denatured heavy meromyosin²⁰ prepared from adult rabbit myosin as described previously²¹. When tested in complement fixation assays²¹, this antiserum gave one peak of reactivity using adult fast myosin from either rabbits or rats as test antigen. Antibodies cross-reacting with slow and neonatal myosin were removed by absorption against myosins prepared from the adult rat soleus muscle and from muscle of 8-day old rats. The myosins and antiserum were mixed in 0.4 M NaCl, 30 mM Tris-HCl (pH 7.4) and incubated for 1 h at 4 °C. The mixture was then dialysed at the same temperature against 30 mM NaCl, 20 mM Tris-HCl (pH 7.4) to precipitate the myosin and myosin-antibody complexes⁸. The specificity of the antibody remaining in the supernatant was verified by indirect immunofluorescence; the absorbed anti-fast serum did not react with fibres containing slow myosin in adult rat muscles nor did it react with neonatal muscle. The absorbed serum retained reactivity with adult fast fibres⁸. An antiserum to neonatal myosin was prepared by injecting rabbits with SDS-denatured neonatal heavy meromyosin (from 8-day old rats). The serum was passed over a Sepharose 4B column to which native neonatal myosin had been covalently attached; the bound immunoglobulins were then eluted with 3 M MgCl₂ and the eluate was dialysed against phosphate-buffered saline (PBS) and then against 30% glycerol in PBS. The purified immunoglobulins were absorbed against adult fast myosin, adult slow myosin and embryonic myosin (prepared from cultured L6 cells, see ref. 14) as described above. The specificity of the antibody was again verified by indirect immunofluorescence. The antisera from which both of these antibodies were made recognize the myosin heavy chain subunit as determined by their reaction on nitrocellulose sheets²² after transfer of myosin previously separated by SDS-gel electrophoresis. Immunoelectron microscopy was also used to show that the purified anti-neonatal myosin immunoglobulins react with myosin-containing thick filaments but not with thin filaments in myofibrils prepared from muscles of 7-day old rats (I. Pinset-Härström *et al.*, unpublished observations). Denervation of the lower right hind limb muscles of rats was performed 7–8 days after birth by removing a 4–5-mm piece of the sciatic nerve. At 4 weeks after denervation, electrical stimulation of the sciatic nerve proximal to the point at which it had been severed did not evoke contractions in the muscles of the lower leg. The wet weight of the denervated muscles at this time was about half that of the contralateral leg. The gastrocnemius muscles were removed at various times after denervation, mounted for immunocytochemistry and histochemistry and frozen in isopentane maintained in liquid nitrogen. Cryostat sections 10 µm thick were cut at –20 °C and allowed to dry in air. Incubations with the antibodies to myosin were carried out at 37 °C for 1 h or at 4 °C for 16–20 h. Incubation with the fluorescein-labelled second antibodies (Nordic) was at 37 °C for 30 min. Sections were mounted in Gelvatol (Monsanto). The contralateral innervated muscle of 22-day old operated rats (*a*, *b*), the 2-week-denervated muscle from 22-day old rats (*c*, *d*), and the 4-week-denervated muscle from 35-day old rats (*e*, *f*). The antibody against adult fast myosin was used in *a*, *c* and *e*; the antibody against neonatal myosin was used in *b*, *d* and *f*. For each pair of photos, the same region of two serial sections is shown. It is thus possible to determine the staining of the same fibre with each of the two antibodies. $\times 170$.

determined. Using a specific anti-slow myosin serum, we have found that the gastrocnemius muscle of 1-week old rats does contain regions composed of up to 40–50% slow myosin-containing fibres (results not shown). These fibres are therefore present at the time when the denervation is carried out. In the gastrocnemius muscle that had been denervated for 2 weeks, regions are found in which large-diameter fibres react with the anti-slow serum (Fig. 2*a*). It is thus possible that these slow myosin-containing fibres did not develop in the absence of innervation but rather were those present at the time of denervation. Almost all the fibres in this region of the denervated muscle now also react with fast myosin antibody (Fig. 2*b*), similar to the result shown in Fig. 1*c*. Neonatal myosin immunoglobulins stain all the fibres, although the slow myosin-containing fibres are less intensely stained (Fig. 2*c*). Finally, the embryonic myosin antibody gives a weak but significant staining on some fibres (Fig. 2*d*).

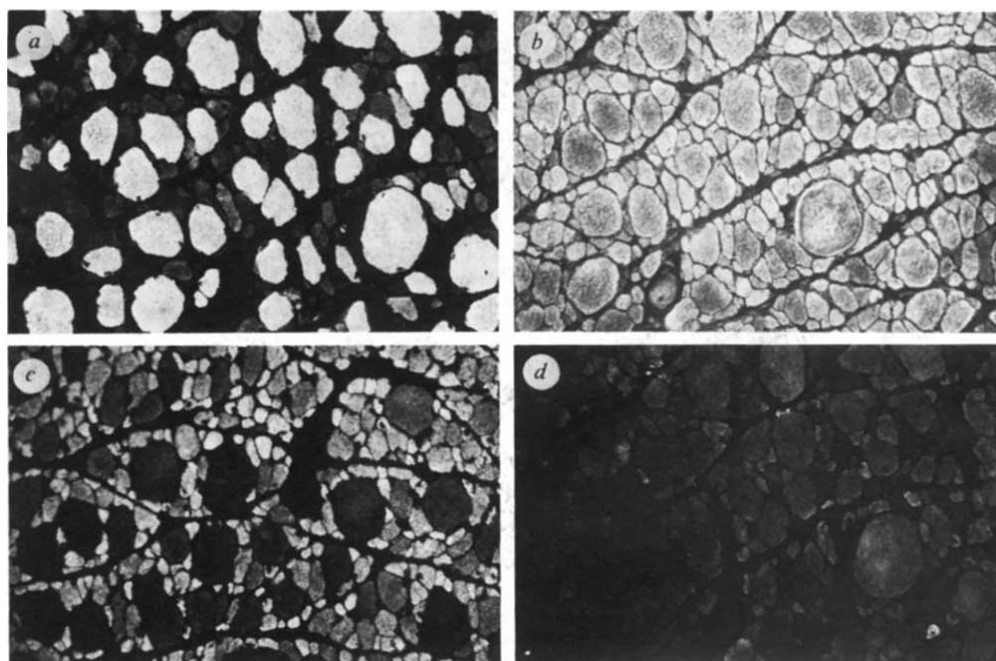
To confirm the appearance of adult fast myosin in the denervated muscles, we analysed the myosin present in extracts of control (unoperated) or denervated muscles by electrophoresis in native conditions^{11,12}. As shown in Fig. 3, the bands corresponding to neonatal myosin⁸ were predominant up to 10 days after birth; three bands were seen at this time, probably because of the isoenzymes formed by the different combinations of light chains, as has been shown for adult fast myosin¹³. By 13 days after birth, fast myosin isoenzymes begin to appear in control and denervated muscles and by 16–22 days the adult fast forms are predominant in both (Fig. 3). At no time does the slow

myosin band represent a major component; it is seen as a minor species in control muscles and some of the denervated ones (Fig. 3). These results indicate that the kinetics of appearance of adult fast myosin are quite similar to those of the control muscles; analysis of myosin isoenzymes in the innervated muscles of operated rats also shows a similar time course (not shown). Bands corresponding to the neonatal myosin isoenzymes are very faint in the 22-day denervated sample (Fig. 3). The staining of the denervated muscles with the neonatal antibody (Figs 1, 2) may therefore indicate that neonatal myosin is present as a minor component.

Polypeptide mapping^{8,14} was also used to examine the heavy chain form found in the denervated muscles. Myosins were prepared from 6-day old control rats and from several denervated and contralateral innervated muscles, and the purified proteins were denatured with SDS. After partial proteolysis in the presence of SDS with one of three proteases (chymotrypsin, *Staphylococcus aureus* protease or subtilisin)^{8,14}, the polypeptides were analysed by SDS-gel electrophoresis (Fig. 4). For each enzyme, neonatal myosin (*a*, *b*, *c*) gave a cleavage pattern different from either the denervated (*a'*, *b'*, *c'*) or contralateral innervated (*a''*, *b''*, *c''*) muscle myosins; the latter two were indistinguishable in this analysis.

All of the results presented here indicate that the myosin found in the gastrocnemius muscle of 3-week old rats, whose sciatic nerve had been sectioned 1 week after birth, is indistinguishable from adult fast myosin by immunological, electrophoretic and protein chemical criteria. The continued

Fig. 2 Indirect immunofluorescence using antibodies against adult fast, adult slow, neonatal and embryonic myosin on cryostat sections of a 2-week-denervated gastrocnemius muscle from 22-day old rats. The antibodies to adult fast and neonatal myosin were those described in Fig. 1 legend. The antibody reacting with adult slow myosin was prepared by injecting guinea pigs with SDS-denatured adult rat cardiac heavy meromyosin^{20,21}. This antiserum cross-reacts strongly with adult rat slow myosin, as determined by microcomplement fixation^{14,21} or by indirect immunofluorescence on control rat muscles containing slow fibres (results not shown). The antiserum was absorbed with neonatal myosin as described in Fig. 1 legend; indirect immunofluorescence showed that the absorbed serum did not react with either control adult fast fibres or with those neonatal fibres reacting with antibody to neonatal myosin. The antibody against embryonic myosin was prepared by injecting rabbits with SDS-denatured myosin heavy chain prepared from L6 myotube cultures¹⁴. Immunoglobulins were purified from this antiserum by affinity chromatography on a column of L6 myosin. When used in indirect immunofluorescence, this immunoglobulin preparation stained fetal muscle strongly but did not react with fibres of neonatal (7 days post-natal) or adult animals. The antisera from which these two antibodies were made also recognize the myosin heavy chain subunit, determined as described in the legend to Fig. 1. This figure shows a region of the denervated gastrocnemius muscle rich in large, hypertrophied fibres. The antibodies used were: *a*, anti-adult slow; *b*, anti-adult fast; *c*, anti-neonatal; *d*, anti-embryonic. All are of the same region in four serial cryostat sections; the staining of the same fibres with each of the four antibodies can thus be evaluated. *a*, *b* and *d* are easily compared; note, however, that the central part of *c* corresponds to the upper part of *a*, *b* and *d*. $\times 170$.



presence of the nerve did not therefore seem to be required for the heavy chain isoenzyme transitions which normally occur during this period. Previous biochemical studies of denervated newborn rat muscle had examined only the myosin light chains^{5,6}. These subunits are the same in newborn fast muscles and the adult muscles⁵, and thus no transition takes place.

Although the kinetics of appearance of fast myosin were not markedly altered by nerve section, denervation did seem to slow down the disappearance of anti-neonatal myosin staining which in innervated muscles has occurred in 20–30% of the fibres by 22 days after birth (Fig. 1*b*). This resulted in a double staining of almost all fibres of the denervated muscle with antibodies to both fast and neonatal myosin at this time (Figs 1*c*, *d*, 2). However, adult fast myosin was the predominant protein accumulated (Figs 3, 4). Neonatal myosin gradually disappeared from most of the denervated fibres, as seen by

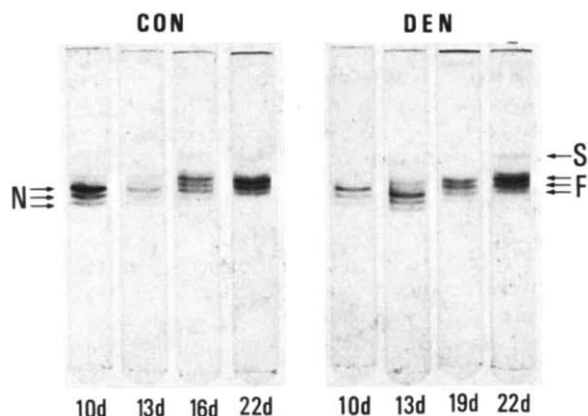


Fig. 3 Electrophoresis of native myosin in non-dissociating conditions. Crude extracts from denervated gastrocnemius muscles (DEN) and from control gastrocnemius muscles from unoperated rats (CON) were analysed by electrophoresis on polyacrylamide gels in the presence of sodium pyrophosphate buffer as previously described⁸. The times indicated are days after birth, denervation having been performed at day 7. Arrows indicate the positions of the bands corresponding to neonatal myosin (N), adult fast myosin (F) and adult slow myosin (S).

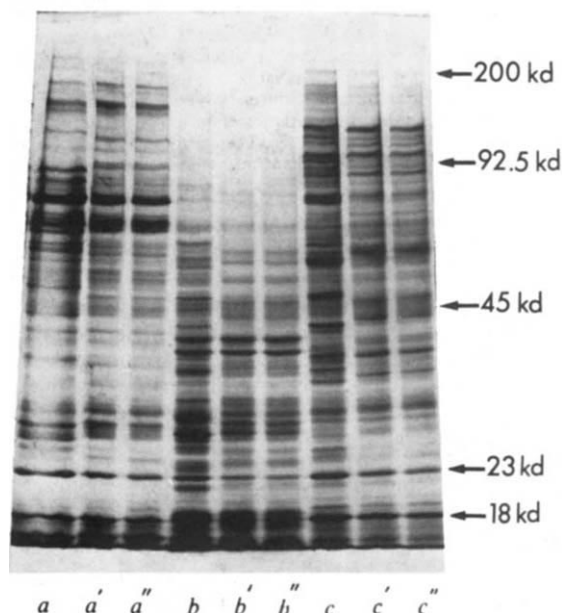


Fig. 4 Polypeptide mapping of the myosin heavy chain. Myosins were prepared by a protocol described previously²¹ from the hind leg muscle of 6-day old rats, and from the denervated and contralateral innervated gastrocnemius muscles of 22-day old rats that had been denervated for 2 weeks. The myosin from the ammonium sulphate fractionation was further purified by passage over a Sepharose 2B column²³. The purest fractions were pooled, placed in dialysis sacks and concentrated against dry Ficoll (Pharmacia). Proteolytic digestion of the myosin was carried out in the presence of 0.5% SDS as described¹⁴ using chymotrypsin (*a*, *a'*, *a''*), *S. aureus* V8 protease (*b*, *b'*, *b''*) or subtilisin (*c*, *c'*, *c''*) at final concentrations of 50 $\mu\text{g ml}^{-1}$, 40 $\mu\text{g ml}^{-1}$ and 0.2 $\mu\text{g ml}^{-1}$ respectively. Lanes *a*, *b*, *c* are myosin prepared from the 6-day old muscles, lanes *a'*, *b'*, *c'* are myosin from the denervated muscles and lanes *a''*, *b''*, *c''* are myosin from the contralateral innervated muscles. Electrophoresis was performed on SDS-polyacrylamide slab gels composed of 12.5% acrylamide and 0.1% bisacrylamide. Molecular mass markers (BioRad) are myosin heavy chain (200,000 (20K)), phosphorylase (92.5K), ovalbumin (45K). The two lowest molecular weight markers are the rat myosin light chains LC1F (23K) and LC2F (18K).

staining at 35 days after birth (Fig. 1e,f). At 35 days the neonatal myosin staining in the denervated muscles is localized to small fibres (Fig. 1f).

It has not been possible to determine, in the experiments shown here, whether development of fibres containing slow myosin is blocked by denervation as others have suggested based on myosin light chain analysis^{5,6}. The fibres staining with the slow myosin antiserum in the denervated gastrocnemius muscle (Fig. 2) were very possibly those slow myosin-staining fibres present at the time of denervation. These observations could explain the contradictory claims concerning slow muscle development after neonatal denervation^{2,5}.

From the results presented here it would appear that the lack of a nerve-imposed activity pattern^{15,16} or the absence of trophic substances¹⁷ secreted by the nerve does not inhibit the accumulation of adult fast myosin in the denervated muscle. This absence of a controlling influence of the nerve might suggest

that initiation of the neonatal myosin → fast myosin transition is an endogenously programmed event in the muscle fibre. It remains possible, however, that the influence already exerted by the nerve by post-natal days 7–8 (when denervation was carried out) is sufficient to initiate the transition. Alternatively, some extrinsic factors could operate in the manner of a hormone. Indeed the level of circulating thyroid hormone seems to influence the myosin types present in adult skeletal¹⁸ and cardiac¹⁹ muscle. Perhaps similar factors affect skeletal muscle development.

We thank Dr Marion Ecob for comments, Dr Pierre Benoit for experimental help, Mrs C. Wisniewsky for technical assistance, and Professor François Gros for support. G.S.B.-B. and L.B.B. were the recipients of Muscular Dystrophy Association postdoctoral fellowships. This work was supported by the CNRS, INSERM and the Muscular Dystrophy Association of America.

Received 26 May; accepted 10 August 1982.

1. Ecob, M. S. *J. neurol. Sci.* (in the press).
2. Engel, W. K. & Karpatis, G. *Dev. Biol.* **17**, 713–723 (1968).
3. Shafiq, S. A., Asiedu, S. A. & Milhorat, A. T. *Expl. Neurol.* **35**, 529–540 (1972).
4. Brown, M. D. *Nature* **244**, 178–179 (1973).
5. Rubinstein, N. A. & Kelly, A. M. *Dev. Biol.* **62**, 473–485 (1978).
6. Ishiura, S., Nonaka, I., Sugita, H. & Mikawa, T. *Expl. Neurol.* **73**, 487–495 (1981).
7. Gauthier, G. F., Lowey, S. & Hobbs, A. W. *Nature* **274**, 25–29 (1978).
8. Whalen, R. G. *et al. Nature* **292**, 805–809 (1981).
9. Whalen, R. G. *Adv. physiol. Sci.* **5**, 63–69 (1981).
10. Rubinstein, N. A. & Kelly, A. M. *J. Cell Biol.* **90**, 128–144 (1981).
11. Hoh, J. F. Y., McGrath, P. A. & White, R. I. *Biochem. J.* **157**, 87–95 (1976).
12. d'Albis, A., Pantaloni, C. & Bechet, J.-J. *Eur. J. Biochem.* **99**, 261–272 (1979).
13. Lowey, S., Benfield, P. A., Silberstein, L. & Lang, L. M. *Nature* **282**, 522–524 (1979).

14. Whalen, R. G., Schwartz, K., Bouveret, P., Sell, S. M. & Gros, F. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5197–5201 (1979).
15. Salmons, S. & Henriksson, J. *Muscle Nerve* **4**, 94–105 (1981).
16. Jolesz, F. & Sreter, F. A. *Rev. Physiol.* **43**, 531–552 (1981).
17. Gutmann, E. A. *Rev. Physiol.* **38**, 177–216 (1976).
18. Johnson, M. A., Mastaglia, F. L., Montgomery, A. G., Pope, B. & Weeds, A. G. *FEBS Lett.* **110**, 230–235 (1980).
19. Hoh, J. F. Y., McGrath, P. A. & Hale, P. T. *J. molec. cell. Cardiol.* **10**, 1053–1076 (1978).
20. Lompré, A.-M., Bouveret, P., Leger, J. & Schwartz, K. *J. immun. Meth.* **28**, 143–148 (1979).
21. Schwartz, K., Lompré, A.-M., Bouveret, P., Wisniewsky, C. & Swynghedauw, B. *Eur. J. Biochem.* **104**, 341–346 (1980).
22. Tobin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
23. Whalen, R. G., Butler-Browne, G. S. & Gros, F. *J. molec. Biol.* **126**, 415–431 (1978).

Preferential effect of γ interferon on the synthesis of HLA antigens and their mRNAs in human cells

David Wallach, Marc Fellous* & Michel Revel

Department of Virology, The Weizmann Institute of Science, Rehovot, Israel

* Unité d'immunogénétique Humaine, Institut Pasteur, 75015 Paris, France

Interferons produce a variety of biological effects on cells. They induce resistance to virus proliferation^{1,2}, inhibit cell growth³, modify cell structure and differentiation^{3–5}, stimulate some immune functions and inhibit others⁶. However, the different interferon (IFN) species may vary in their mechanism of action and, hence, in their relative efficiency for inducing each of the effects. IFN- γ (type II) appears to show stronger immunoregulatory^{7–9} and growth inhibitory effects^{10–12} than antiviral effects¹³, but this conclusion has been challenged in other reports¹⁴. The aim of the present work is to compare the action of IFN- γ and other (type I) interferons on the induction of (2'–5') oligo(A) synthetase which is probably part of the antiviral response^{15,16} and the induction of the histocompatibility HLA-A, -B, -C antigens^{17–19}. We have shown previously that the induction of both proteins is regulated by interferons at the mRNA level^{20,21}, but show here that IFN- γ from stimulated human lymphocytes²² and from monkey cells transfected by cloned human IFN- γ cDNA²³ induced the HLA-A, -B, -C and β_2 -microglobulin mRNAs or proteins at concentrations over 100 times lower than those needed to induce the (2'–5') oligo(A) synthetase and the antiviral state. This difference was not found with IFN- α and - β (type I).

The experiment in Fig. 1a and b compares the effects of IFN- β and IFN- γ on the amounts of cell surface HLA, on the (2'–5') oligo(A) synthetase and on the inhibition of vesicular stomatitis virus (VSV) replication measured simultaneously on the same human fibroblastoid SV80 cells. As the activities of interferons vary from one cell line to another, it was essential to first compare the antiviral potency of IFN- β and IFN- γ on the same cells. This was done by using a radioimmunoassay for

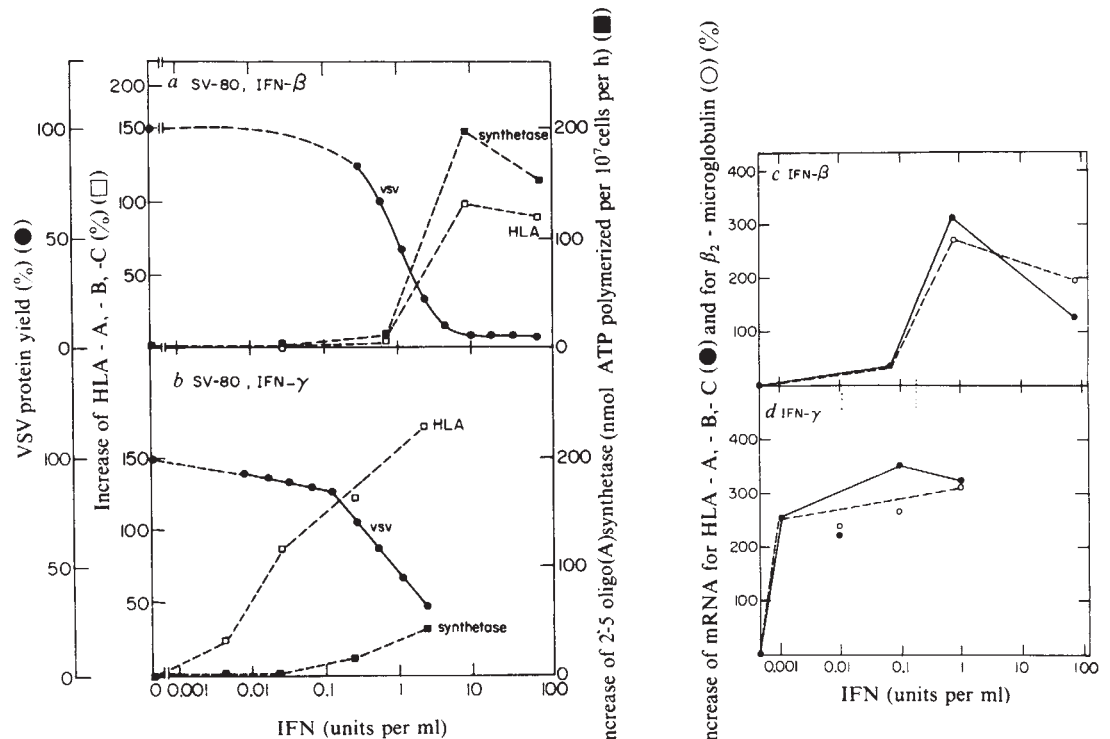
VSV proteins that we have developed for the accurate quantitation of interferon's antiviral effect^{24,25}. The concentration of each IFN required to reduce the yield of VSV proteins on the SV80 cells by 50% was determined (Fig. 1a, b) and is by definition 1 unit per ml on SV80 cells. This corresponds to about 10 units per ml of IFN- β titrated on human diploid fibroblasts, and 100 units per ml of IFN- γ titrated on Wish cells. The (2'–5') oligo(A) synthetase was measured by mixing cytoplasmic extracts of the treated cells with poly(rI)·(rC)-agarose beads, incubating the beads with [α -³²P]ATP and purifying the (2'–5') oligo(A) synthetase formed on alumina columns²⁶. The cell surface HLA was quantitated by the binding

Table 1 Comparison of the effect of α , β and γ interferons on HLA proteins and (2'–5') oligo(A) synthetase in Ramos cells

IFN	U ml ⁻¹	(2'–5') oligo(A) synthetase (nmol per 10 ⁷ cells per h)	HLA-DR c.p.m.	HLA-A, -B, -C c.p.m.	Increase (%)
None		6.6	5,000	5,000	
α	100 (on MDBK)	37.1	–	9,200	84
	400 (on MDBK)	52.0	5,100	11,100	122
β	50 (on FS-11)	19.2	–	8,700	74
	200 (on FS-11)	34.3	5,400	9,500	90
γ	40 (on WISH)	7.5	–	8,900	78
	200 (on WISH)	8.3	5,000	10,300	105

The cells used were lymphoblastoid cell line Ramos grown as before²¹ to 2×10^6 cells per ml in RPMI 1640 (10% fetal calf serum, FCS) and treated with each IFN species for 16 h at 37°C. The cells were washed and counted, and the activity of the (2'–5') oligo(A) synthetase as well as the amount of cell-surface HLA-A, -B, -C and HLA-DR antigens were measured as in Fig. 1. The IFNs were prepared as follows: (1) leukocyte IFN- α subspecies $\beta 3^{35}$, also designated IFN- α B³⁶ was purified to electrophoretic homogeneity by HPLC³⁵ starting from IFN induced by Sendai virus in human myelogenous leukaemia cells (Institut Merieux). This purified IFN- α (a gift from Dr M. Rubinstein) titred 2×10^8 units per mg protein on bovine MDBK cells and 5×10^8 units per mg on human FS-11 fibroblasts. (2) Fibroblast IFN- β , was purified by a modification of the Cibacron blue Sepharose procedure³⁷ from IFN superinduced by poly(rI)·(rC) in human foreskin fibroblasts FS-11²⁴ (InterYeda). Its titre was over 5×10^7 units per mg on FS-11 cells. (3) IFN- γ was prepared as described^{9,22,31} from human peripheral blood mononuclear cells isolated on Ficoll-Hypaque (Pharmacia) and incubated for 48 h with concanavalin A (Con A, $30 \mu\text{g ml}^{-1}$) and phorbol myristate acetate (5 ng ml^{-1}) in RPMI 1640–2% FCS. The IFN- γ was partially purified on controlled pore glass²² and its antiviral activity (10^5 units per mg) was measured by inhibition of VSV growth in Wish cells. The activities of IFN- α and β were also determined with VSV by comparison to International Reference Standards.

Fig. 1 Comparison of the effects of IFN- β and IFN- γ on HLA-A,-B,-C antigens and mRNAs and on the (2'-5')oligo(A) synthetase in SV80 cells. *a, b*, The human fibroblastoid cell line SV80 was grown as before²¹ in Dulbecco's modified Eagle's medium with 10% fetal calf serum (FCS) at 37°C. Confluent monolayers were treated for 16 h with a large range of concentrations of IFN- β (*a*) or IFN- γ (*b*) prepared as in Table 1. The antiviral effect of each IFN on SV80 was measured by infecting one set of cultures with VSV (10 PFU per cell) and measuring after 18 h the yield of VSV proteins in the lysed cells by a radioimmunoassay using rabbit antibodies against purified VSV, followed by ¹²⁵I-protein A as described^{24,25}. The VSV protein yield is shown as per cent of that in cells not treated by IFN (6,000 c.p.m.). One unit per ml of each IFN was defined by the concentration reducing the VSV yield by 50%. In another set of SV80 cultures 16 h after IFN addition, the cells were briefly trypsinized, washed twice with medium-10% FCS and the cell-surface antigens were quantitated as before²¹. In short, 2.5×10^5 cells were washed with RPMI 1640, 0.5% bovine serum albumin, 0.1% Na azide and mixed in this medium with saturating amounts (50 μ l) of mouse monoclonal antibodies W6/32 specific for HLA-A,-B,-C, or M18 specific for β_2 -microglobulin, or L116 specific for HLA-DR. After 1 h at 4°C, the cells were washed as above, incubated for 1 h at 4°C with ¹²⁵I-protein A, washed and counted. The per cent increase in HLA-A,-B,-C is shown relative to the level in cells not treated by IFN (11,500 c.p.m.). For the (2'-5')oligo(A) synthetase assay, 2×10^7 cells were mixed with 1 ml Nonidet P40 extraction buffer²⁶ and 10 μ l of the extracts were used to measure the enzyme activity as detailed before^{26,38}. The synthetase activity is expressed as ³²P-(2'-5') (Ap)nA formed per h and 10^7 cells. Extracts of SV80 cells not treated by IFN gave 0.1 nmol per 10^7 cells per h. *c, d*, SV80 cells were exposed to the same concentrations of (c) IFN- β and (d) IFN- γ as in *a* and *b*, but for 10 h. Poly(A)⁺RNA was prepared from 6×10^8 cells by the LiCl procedure as described²¹. After agarose gel electrophoresis and blotting onto nitrocellulose, the specific mRNAs were quantitated by hybridization to ³²P-labelled cloned cDNA probes as reported in detail before²¹. The HLA probe used here was a HLA-B cDNA clone given by Dr S. M. Weissman²⁹, while the β_2 -microglobulin cDNA clone was a gift from Dr K. Itakura³⁰. Quantitation of the radioactivity hybridized to the mRNA bands was done by spectrophotometric scanning of the autoradiographs. The per cent increase in HLA-A,-B,-C (●) and β_2 -microglobulin (○) mRNAs is shown relative to the level of these mRNAs in untreated cells.



of monoclonal antibodies recognizing a monomorphic determinant of HLA-A,-B,-C²⁷ followed by binding of ¹²⁵I-protein A (ref. 21). Figure 1*a* shows that synthetase and HLA-A,-B,-C were both induced at concentrations of IFN- β which produce the antiviral effect. In contrast, Fig. 1*b* shows that IFN- γ increased the amounts of cell surface HLA-A,-B,-C already at concentrations 100 times lower than those required for the antiviral state. The same was observed for β_2 -microglobulin (not shown). Induction of (2'-5')oligo(A) synthetase by IFN- γ could, however, be detected only in the range of concentrations giving the antiviral effect.

Similar observations were made in other cell types such as the lymphoblastoid line Ramos. Table 1 shows that in these cells, IFN- α and IFN- β both produced a dose-dependent increase in the (2'-5')oligo(A) synthetase as well as in cell-surface HLA-A,-B,-C. Again, IFN- γ increased the HLA-A,-B,-C antigens in conditions where synthetase induction was barely detectable. In all cases, cell-surface HLA-DR remained unchanged as reported^{19,28}. The differential effects of IFN- γ do not depend on the actual titres of IFN used but on the comparison between the relative concentrations at which these effects occur. In Wish cells, for example, which are most sensitive to the antiviral effect of IFN- γ , the concentration of IFN- γ needed to increase the cell-surface HLA antigens was again 30 times lower than that needed to inhibit VSV growth (not shown).

We have shown²¹ that the increase in cell-surface HLA-A,-B,-C induced by type I IFNs is preceded by an even larger increase in the amount of cytoplasmic mRNA coding for these proteins. We therefore compared the effects of IFN- γ and IFN- β on the levels of HLA mRNAs in SV80 cells as described in Fig. 1 legend. The mRNAs were quantitated by hybridization of electrophoretic blots of total cell poly(A)⁺ RNA to a cloned

HLA-B cDNA probe which detects all three HLA-A,-B,-C mRNAs²⁹ and to a cloned β_2 -microglobulin cDNA probe³⁰. Figure 1*c* and *d* shows that IFN- γ induces already maximum (fourfold) increases in HLA and β_2 -microglobulin mRNAs at much lower concentrations than does IFN- β . By comparison with Fig. 1*b*, it can be seen that IFN- γ diluted 1,000 times more than the concentration needed to inhibit VSV growth, is able to fully induce the HLA antigen mRNAs. At these low concentrations, IFN- γ did not induce the (2'-5')oligo(A) synthetase mRNA (not shown).

It is very likely that the increase in HLA mRNAs is responsible for the increase in the antigens. To measure directly the rate of HLA protein synthesis, we labelled the SV80 cells with ³⁵S-methionine for 1 h, lysed the cells and immunoprecipitated the β_2 -microglobulin together with their associated HLA-A,-B,-C chains using polyclonal antibodies against β_2 -microglobulin. The proteins were separated by gel electrophoresis and the labelling of HLA-A,-B,-C and β_2 -microglobulin chains was quantitated. Figure 2 shows that IFN- β and IFN- γ induced a marked 4- to 5-fold increase in the rate of synthesis of the HLA proteins, but with the same difference in the amount of each IFN needed for maximum induction as that seen in Fig. 1 for cell-surface HLA. The stimulation of HLA synthesis by IFN- β required concentrations sufficient to induce the antiviral state (Fig. 2*a*), while IFN- γ stimulated HLA synthesis at much lower concentrations (Fig. 2*b*). A comparison of Figs 1 and 2 also shows that the synthesis of HLA proteins is, like their mRNAs²¹, stimulated by IFNs to a higher extent than the cell-surface HLA antigens.

The striking dissociation between IFN- γ 's antiviral effect on VSV and the (2'-5')oligo(A) synthetase induction on the one hand, and on HLA-A,-B,-C and β_2 -microglobulin induction

on the other hand, raises the question of whether or not it is the same molecule of IFN- γ which produces both effects. As partially purified IFN- γ preparations from stimulated lymphocytes could contain lymphokines or other IFNs¹⁴, we examined this possibility by inducing IFN- γ by two different methods in human white blood cells. Figure 3a and b show that with both preparations the HLA-inducing and the antiviral activity clearly co-purify on Ultrogel AcA 54 or on Sephacryl S-200 and elute as a single peak, separated from the bulk of proteins and corresponding to the characteristic high molecular weight position of IFN- γ ³¹. No HLA-inducing factor other than IFNs is known, and none could be detected in experiments such as that of Fig. 3.

Recently, the cDNA which codes for human IFN- γ has been cloned²³. Monkey kidney fibroblasts of the COS-7 cell line,

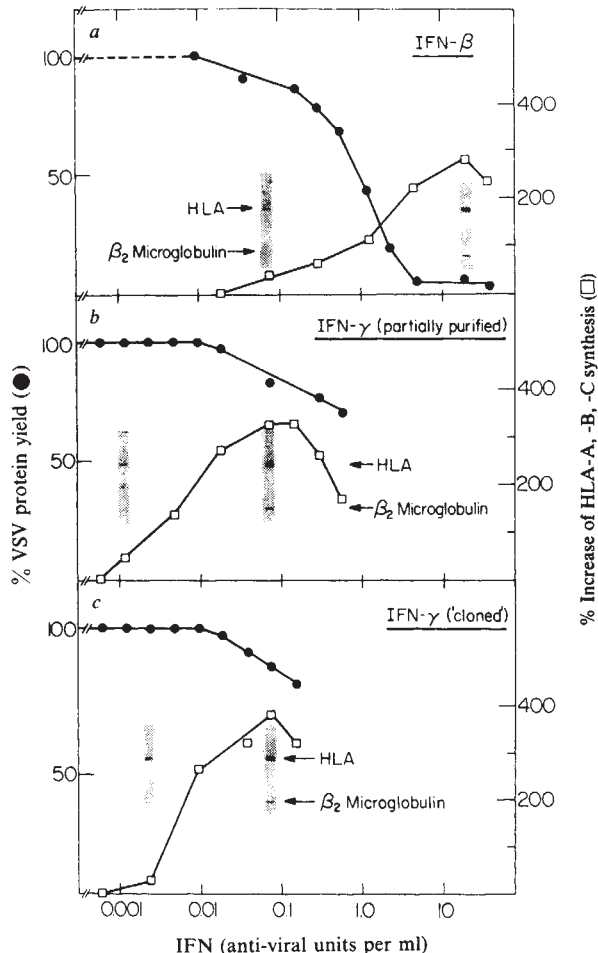


Fig. 2 Comparison of the effects of IFN- β and IFN- γ on the synthesis of HLA-A, -B, -C proteins in SV80 cells. Cultures of SV80 cells were treated for 18 h as in Fig. 1 with (a) IFN- β and (b) IFN- γ prepared as in Table 1 or (c) IFN- γ produced in COS-7 monkey cells transfected by the pSV-Y-69 plasmid DNA containing the human IFN- γ cDNA under the control of the SV40 late promoter²³. The antiviral effect of each IFN on VSV growth was measured and compared on SV80 cells as in Fig. 1 (●). To measure the rate of HLA protein synthesis, the cells were washed with methionine-free medium and incubated for 1 h with 0.5 mCi ml⁻¹ ³⁵S-methionine (800 Ci mmol⁻¹; Amersham) in this medium-2% FCS. The cells were washed in cold phosphate-buffered saline, and extracted with 0.5% Nonidet P40 in 0.15 M NaCl, 10 mM Tris-HCl pH 7.4, 5 mM MgCl₂, 1 mM phenylmethylsulphonyl fluoride (PMSF), 1 mM 1-tosylamide 2-phenylethyl chloromethyl ketone (TPCK). After centrifugation, the cytoplasmic extracts were incubated for 1 h at 4°C with polyclonal antibodies against β_2 -microglobulin (DAKO) and with 20 mM EDTA. Protein A-Sepharose CL-4B beads (Pharmacia) were added for 1 h at 4°C with stirring. The beads were washed repeatedly in 0.15 M NaCl, 10 mM Tris-HCl pH 7.5, 5 mM EDTA, 0.5% Nonidet P40, 0.5 mg bovine serum albumin and the bound proteins were solubilized in dodecyl sulphate before electrophoresis on a 10-20% polyacrylamide gradient gel. After autoradiography, the radioactivity in the HLA-A, -B, -C and β_2 -microglobulin bands was quantitated by scanning. The per cent increase in the HLA-A, -B, -C band is shown (□) relative to the level in cells not treated by IFN. Examples of the autoradiography are shown for a few indicated IFN concentrations.

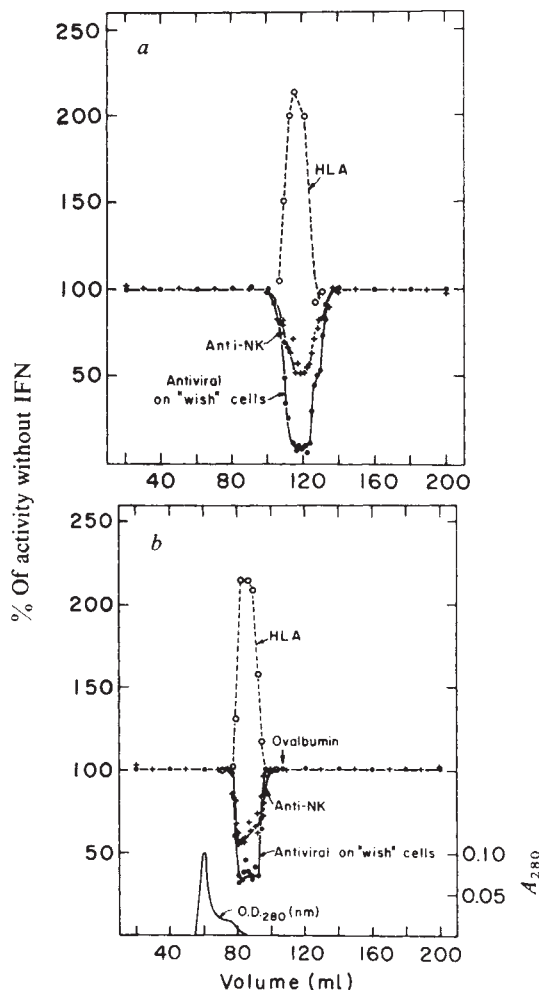


Fig. 3 Co-purification of the HLA-inducing, anti-NK and antiviral activities of IFN- γ . Two preparations of IFN- γ were induced as described in Table 1, with either (a) 3 μ g ml⁻¹ phytohemagglutinin (PL Biochemicals) or (b) 30 μ g ml⁻¹ Con A in the presence of 5 ng ml⁻¹ phorbol myristate acetate. After adsorption and elution from controlled pore glass²², samples of 0.7 ml were applied to a 1.7 $\times$ 110 cm column of either Ultrogel AcA 54 (LKB) (a) or Sephacryl S-200 (b) (Pharmacia), and eluted with 1 M NaCl, 20 mM Na phosphate buffer pH 7.4 and ethylene glycol (25% in a and 40% in b). Fractions of 1 ml were collected in 1% FCS and used at a final dilution of 1:40 to measure by radioimmunoassay²⁴ the inhibition of VSV growth in Wish cells (antiviral effect) and the HLA-A, -B, -C inducing activity in SV80 cells (as in Fig. 1). The anti-NK effect was measured as before^{9,34} by ⁵¹Cr release from prelabelled HeLa cells treated by the IFN- γ fractions and incubated for 3 h with non-adherent human peripheral blood leukocytes which had been pretreated for 18 h with 500 units per ml IFN- β . The leukocyte to HeLa cell ratio was 40:1. Without IFN- γ , 34% of the HeLa cells were killed in a and 60% in b. The three IFN- γ activities are shown as per cent decreases in VSV proteins (●), increase in HLA antigens (□) or decrease in specific chromium release (+) relative to controls without IFN. Ovalbumin (molecular weight 44,000) was used as marker in b.

transfected by the recombinant pSV-Y-69 DNA produce mature IFN- γ ²³, which is free of the contaminants found in IFN- γ prepared from human blood cells. When tested for the stimulation of HLA synthesis, as shown in Fig. 2c, this IFN- γ was identical to the partially purified human blood cell IFN- γ used in Fig. 2b. A strong stimulation of HLA synthesis in SV80 cells was seen at concentrations much lower than those needed to produce the antiviral effect, in sharp contrast to the behaviour of fibroblast IFN- β (Fig. 2a). The more efficient HLA-inducing activity was, therefore, transferred to the monkey cells by the IFN- γ cDNA, demonstrating that this is a genuine property of the recently identified IFN- γ molecule, whose amino acid sequence is very different from that of type I IFNs (ref. 23).

With regard to the biological function of the preferential effect of IFN- γ on HLA antigen synthesis in non-immune cells, IFN- γ might enhance the presentation of viral antigens on the surface of virus-infected cells thereby increasing their destruc-

tion by cytotoxic lymphocytes³². For the same antiviral activity, IFN- γ is also much more active than type I IFNs in making human cells resistant to natural killer (NK) cells⁹. This anti-NK effect may be important in protecting some cells against indiscriminate killing by NK cells³³. By analogy with HLA induction, this anti-NK effect of IFN- γ was seen in SV80 and HeLa cells at concentrations much lower than those needed to inhibit VSV growth^{9,34}. Figure 3 shows that this anti-NK activity elutes together with the HLA-inducing and antiviral activities of IFN- γ . Moreover, the IFN- γ produced by the cloned cDNA also induced the anti-NK effect at very low concentrations (not shown). At least one immunoregulatory function is, therefore, induced by IFN- γ in non-immune cells with the same efficiency as the HLA antigens are induced, and much more efficiently than the (2'-5')oligo(A) synthetase or the antiviral state⁹.

These results confirm that IFN- γ and type I IFNs share the ability to induce (2'-5')oligo(A) synthetase and HLA proteins but in the case of IFN- γ there is a dissociation between the two effects and HLA synthesis is stimulated at concentrations much lower than those needed for the synthetase induction. Such a difference may explain some of the biological effects of IFN- γ , although there are probably other cellular proteins which could be differentially induced by type I and type II IFNs¹². The differential effect of IFN- γ may result from the presence of cell-surface receptors distinct from those of type I IFNs³⁹. Many cells are poorly sensitive to the antiviral effect of IFN- γ ¹⁴, and the induction of HLA synthesis could be a more sensitive assay for IFN- γ in such cells.

We thank Dr S. M. Weissman and K. Itakura for gifts of the HLA and β_2 -microglobulin cDNA clones, Dr M. Rubinstein and S. Cimbalista for gifts of IFNs, and Dr P. Gray and D. V. Goeddel (Genentech) for the gift of IFN- γ produced by cloned cDNA. We thank R. Schluskel, U. Nir, G. Merlin and F. Rosa for their collaboration and Ms S. Budilovsky for technical assistance. This work was supported in part by a grant from NCRD (Israel) and GSF (München, Germany).

Received 8 March; accepted 26 August 1982.

- Isaacs, A. & Lindenmann, J. *Proc. R. Soc. B* **147**, 258-267 (1957).
- Friedman, R. M. & Chang, E. H. in *Interferons and Their Actions* (ed. Stewart, W. E. II) 145-152 (CRC Press, Ohio, 1977).
- Gresser, I. & Tovey, M. G. *Biochim. biophys. Acta* **516**, 231-247 (1978).
- Pfeffer, L. M., Wang, E. & Tamm, I. *J. cell. Biol.* **85**, 9-17 (1980).
- Luftig, R. B. *et al. J. Virol.* **23**, 799-810 (1977).
- De Maeyer, E. in *The Biology of the Interferon System* (eds De Maeyer, E., Galasso, G. & Schellekens, H.) 203-209 (Elsevier, Amsterdam, 1981).
- Sonnenfeld, G., Mandel, A. D. & Merigan, T. C. *Cell. Immun.* **34**, 193-206 (1977).
- Virelizier, J. L., Chan, E. L. & Allison, A. C. *Clin. exp. Immun.* **30**, 299-304 (1977).
- Wallach, D. in *Human Lymphokines, Biological Immune Response Modifiers* (eds Khan, A. & Hill, N. O.) 435-449 (Academic, New York).
- Salvin, S. B., Younger, J. S., Nishio, J. & Netar, R. *J. natn. Cancer Inst.* **55**, 1233-1235 (1975).
- Crane, J. L., Glasgow, L. A., Kern, E. & Younger, J. S. *J. natn. Cancer Inst.* **61**, 871-874 (1978).
- Rubin, B. Y. & Gupta, S. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5928-5932 (1980).
- Dianzani, F., Salter, L., Fleischmann, W. R. Jr & Zucca, M. *Proc. Soc. exp. Biol. Med.* **159**, 94-97 (1978).
- Epstein, L. B. in *Interferon 3* (ed. Gresser, I.) 13-44 (Academic, New York, 1981).
- Knight, M. *et al. Nature* **288**, 189-191 (1980).
- Revel, M. in *Interferon 1* (ed. Gresser, I.) 101-163 (Academic, New York, 1979).
- Lindahl, P., Leary, P. & Gresser, I. *Eur. J. Immun.* **4**, 779-794 (1974).
- Heron, I., Hokland, M. & Berg, K. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6215-6219 (1978).
- Fellous, M., Kamoun, M., Gresser, I. & Bono, R. *Eur. J. Immun.* **9**, 446-452 (1979).
- Shulman, L. & Revel, M. *Nature* **287**, 98-100 (1980).
- Fellous, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 3082-3086 (1982).
- Yip, Y. K., Pang, R. H. L., Urban, C. & Vilcek, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1601-1605 (1981).
- Gray, P. W. *et al. Nature* **295**, 503-508 (1982).
- Weissenbach, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7152-7156 (1980).
- Wallach, D. *et al. J. Interferon Res.* (submitted).
- Merlin, G., Revel, M. & Wallach, D. *Analyst. Biochem.* **110**, 190-196 (1981).
- Barnstable, C. J. *et al. Cell* **14**, 9-31 (1978).
- Vignaux, F. & Gresser, I. *J. Immun.* **118**, 721-723 (1977).
- Sood, A. K., Pereira, D. & Weismann, S. M. *Proc. natn. Acad. Sci. U.S.A.* **78**, 616-620 (1981).
- Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Itakura, K. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6613-6617 (1981).
- Vilcek, J. *et al. Miami Winter Symp.* **18**, 331-340 (1981).
- Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 96 (1979).
- Trichieri, G. & Santoli, D. *J. exp. Med.* **147**, 1314-1333 (1978).
- Wallach, D. *J. Interferon Res.* **2** (in the press).
- Rubinstein, M. *et al. Archs Biochem. Biophys.* **210**, 307-318 (1981).
- Goeddel, D. V. *et al. Nature* **290**, 20-26 (1981).
- Knight, E. Jr & Fahey, D. *J. biol. Chem.* **256**, 3609-3611 (1981).
- Schattner, A. *et al. Lancet* **ii**, 497-500 (1981).
- Branca, A. A. & Baglioni, C. *Nature* **294**, 768-770 (1981).

Direct evidence of homology between human DC-1 antigen and murine I-A molecules

M. Rosa Bono & Jack L. Strominger

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

The nature of the interactions between cells of the immune system has been shown to be directly related to a group of antigens present on the surface of B lymphocytes¹. These antigens are under genetic control of the HLA-D/DR region of the human major histocompatibility complex (MHC) located on chromosome 6². DR antigens (the products of the *HLA-DR* locus within this region), have been defined by serological studies. Recently another antigen, DC-1, has been serologically defined^{3,4}. In the mouse, the I region, analogous to the HLA-D/DR region in man, has been subdivided into at least five subregions (A, B, J, E and C) and specific functions and cell populations associated with each of the subregions⁵. Although in man the HLA-D/DR region cannot at present be subdivided, limited NH₂-terminal amino acid sequence data have shown that the DR molecule is homologous to the I-E molecule of the mouse^{6,7}. Furthermore, evidence that DC-1 is not homologous to I-E (and might therefore be homologous to I-A) has been reported⁸. We now report that, using immunoadsorbents prepared with monoclonal antibodies specific for DR or DC-1 antigens, these two human alloantigens were purified from a human lymphoblastoid cell line (LB) and the chains were separated. The NH₂-terminal amino acid sequence of the heavy chain of DC-1 antigen shows great structural homology with the murine I-A molecules, thus providing direct evidence of homology of DC-1 to murine I-A.

DR antigens, like the corresponding murine I region encoded antigens, are carried on molecules composed of two glycoprotein chains (α or heavy chain and β or light chain) of apparent molecular weights 34,000 and 29,000. The response to a certain antigen or susceptibility to certain diseases have been related to the expression of a particular allele of DR and I region encoded antigens⁹. Ten alleles of the *HLA-DR* locus have been identified by serological studies¹⁰. DC 1 (identical to MB 1, MT 1 or LB 12) has been identified as having the same cell distribution and a similar molecular structure to the DR antigen. It appears to be immunogenetically associated with the specificities DR 1, 2 and w6 (ref. 3). However, the DC-1 antigen is expressed on a separate molecule from the

Table 1 Amino acid analysis of the α - and β -chains of DR and DC-1 (values in mol per cent)

	α -DR	α -DC-1	β -DR	β -DC-1
Asp	10.0	9.2	10.9	9.3
Thr	4.9	5.0	7.5	7.0
Ser	10.3	9.2	10.8	8.3
Glu	10.4	11.3	15.7	13.1
Pro	4.9	5.5	5.8	6.4
Gly	16.4	14.8	15.4	10.1
Ala	6.9	7.4	7.1	6.6
Val	4.9	4.8	7.3	6.2
Met	1.0	0.7	0.8	0.6
Ile	2.8	2.7	2.6	3.0
Leu	5.6	6.8	7.8	7.3
Tyr	2.0	2.5	3.1	2.9
Phe	3.7	2.8	4.1	3.0
Lys	5.4	5.7	4.5	4.2
His	6.0	7.0	4.4	4.6
Arg	5.1	5.3	7.7	7.8

Values are the average of two separate preparations.

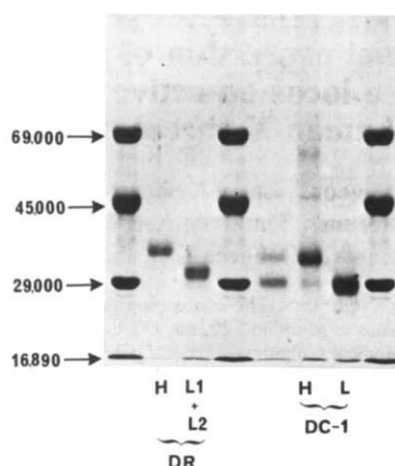


Fig. 1 SDS-polyacrylamide (12%) gel electrophoresis of purified DR and DC-1 antigens. The antigens were obtained from a homozygous B lymphoblastoid cell line, LB, typed as DRw6, w6 and DC-1-positive. Detergent-solubilized membrane proteins were prepared by incubation of frozen cells for 30 min at 4 °C in a buffer solution containing 2% Nonidet P-40, 0.01 M Tris pH 7.8, 0.14 M NaCl, 1 mM MgCl₂ and 0.1 mM PMSF (1 g cells per 4 ml lysis solution). The lysate was centrifuged at 150,000g for 1 h, and the supernatant recovered and centrifuged for an additional 40 min at 12,000g. Before purification of specific antigens, the bulk of actin contaminant was removed by passing the lysate through a Sepharose-4B CL column to which normal rabbit serum was coupled. Purification of DR and DC-1 antigens was done by using immunoabsorbents prepared with a monoclonal antibody L243 to purify DR antigen (obtained from Dr L. Lampson¹³) and Genox 3.53 to obtain the DC-1 antigen (from Dr F. Brodsky¹⁴). Immunoabsorbents were prepared by the method of March *et al.*¹⁹ to give a concentration of 3 mg of protein coupled to 1 ml of Sepharose, then stabilized by cross-linking with a solution of 0.1% glutaraldehyde²⁰. The best yields were obtained when the lysates were first depleted of DR antigens, before purification of DC-1 antigen (yield increased by ~20%). The whole process was done at 4 °C. Lysate corresponding to 20 g of cells was cycled through a 20 ml column of either L243 or Genox 3.53 immunoabsorbent for at least 24 h at a flow rate of 5 ml h⁻¹. Columns were washed with no more than five column bed volumes of a buffer containing 0.14 NaCl, 0.1% DOC and 0.01 M Tris pH 8.0. Elution of the columns was carried out using a buffer of 0.5% DOC, 5% glycerol and 0.1 M Tris pH 11.5. Analysis of the eluted fraction was done by SDS-PAGE. Fractions containing either DR or DC-1 antigen were prepared for amino acid sequencing (Table 2) by extensive dialysis against 0.1% SDS, 1 M Tris pH 8.0 solution, lyophilization, acetone precipitation (5 volumes) and finally resuspension in a 0.1% SDS solution.

DR antigens^{3,4}. Immunogenetic studies¹¹ also suggest that MB 2 (associated with DR 3 and 7) and MB 3 (associated with DR 4, 5 and 8) are alleles of DC-1 and recently MB 2 has also been shown to be on a molecule distinct from DR 7 and similar to the DC-1 (MB 1) antigen¹².

Detergent-solubilized membranes of the LB cell line (DRw6, w6; DC-1) were passed over an immunoabsorbent column prepared with the DR-specific monoclonal antibody L243 (ref. 13). The bound material was eluted at pH 11.5 and subjected to SDS-polyacrylamide gel electrophoresis (PAGE) (Fig. 1). The unbound material was passed over a second column prepared with the DC-1-specific monoclonal antibody Genox 3.53 (refs 4, 14). The bound material was again eluted at high pH and subjected to SDS-PAGE (Fig. 1).

The DR molecule is composed of heavy chain (H) of apparent molecular weight (MW) 34,000 and two light chains of MWs 29,000 (L1) and 28,000 (L2). The DC-1 molecule contains a heavy chain (H) of MW 32,000 and a light chain (L) of MW 27,000. Note also that the light chain of DC-1 has a larger gel mobility than the L2 chain of DR, from which it can also be distinguished by other methods¹². The best yields from 20 g of LB cells were approximately 3 mg of DR antigens and 1 mg of DC-1 antigen.

The individual chains were then separated by hydroxylapatite chromatography (Fig. 2) to study their amino acid compositions and sequences. With this procedure the heavy chain (H) of the DR antigen was separated from the two light chains (L1 + L2) and the H chain of DC-1 was also separated from its L chain. Amino acid analyses showed significant differences between all four of these preparations (Table 1).

The NH₂-terminal amino acid sequences of the α -chains (H) of DR and DC-1 are shown in Table 2. For comparison, data available for the α -chain of the I-A and I-E molecules of the mouse^{1,6,7,12,15} are also shown. Homologous residues between the α -chains of DR and I-E molecules and between the α -chains of DC-1 and the I-A molecules have been underlined. Only 12 residues are available for comparison in the α -chain of I-E. Ten of these are identical in the DR molecule. These data agree with previous results indicating homology between these two antigens^{6,7}. Only eight positions are available for comparison in the α -chain of I-A. Five of these are identical in the DC-1 α -chain if a gap is inserted at the beginning of the DC-1 sequence. A similar gap was introduced to compare the sequence of the α -chain of the I-A molecule of the marmoset to that of the mouse¹⁶. In addition, conservative substitutions are present at two of the three additional positions compared and two additional valine residues are homologous to valines in the marmoset I-A α -chain sequence. The monoclonal antibody which recognizes the marmoset I-A molecule was also shown to react with an HLA-DR-like antigen in a human cell line, although no direct chemical information about this material was reported¹⁶. Four attempts to obtain an NH₂-terminal amino acid sequence of the β -chain of DC-1 yielded no sequence information, suggesting that this chain is blocked at its NH₂-terminus. Treatment with the enzyme which removes pyrrolidone carboxylic acid did not unblock the sequence. Attempts to obtain sequence information by other methods are in progress.

These data are direct evidence that the α -chain of DC-1 is the molecule analogous to the α -chain of the I-A molecule of the mouse. Although DC-1 and DR have similar molecular weights and nearly identical isoelectric points (in contrast to

Table 2 NH₂-terminal amino acid sequence comparisons between the heavy chains of murine I-A and I-E and human DR and DC-1 molecules

Chain	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
DR α	I	K	E	E	H	V	I	I	Q	A	E	F	Y	L	N	P	D	Q	S	G	E	F	M	F	D	F	D	G
I-E α	I	K			H	T	I	I		A		F	Y	L	L	P		R			F							
DC-1 α	E	D	I	V	A	D	-	V	A	Q	L	G	V	N	L	Y	Q	F	Y	L/G	P	-	G	Q	Y/F	-	-	E
I-A α			I	m	A			V		V	Y		*		V	Y												

The NH₂-terminal sequences of the DR and DC-1 α -chains were obtained on a Beckman 890-C sequenator with a Quadrol program. The thiazoline derivatives from each sequence step were converted to phenylthiohydantoin (PTH) amino acids and analysed by HPLC. All of the blanks in the DC-1 α chain are either Ser, Arg or His. The murine sequences are from the references cited in the text. Homologous residues are shown in bold type.

* V residues are found at these two positions in the marmoset I-A α sequence.

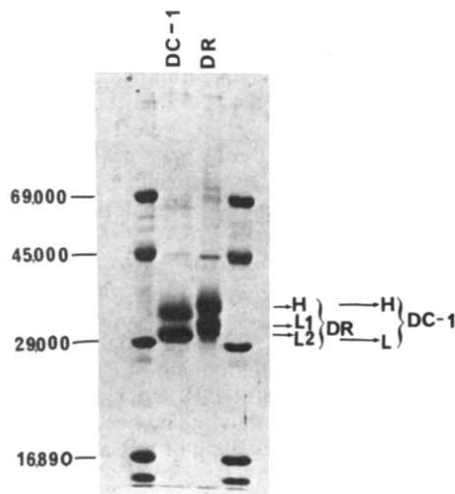


Fig. 2 Polyacrylamide gel electrophoresis of the separate α - and β -chains of DR and DC-1 antigens. Separation was achieved by chromatography on a hydroxylapatite column in the presence of 0.1% SDS and 1 mM dithiothreitol, using a linear phosphate gradient between 0.1 mM and 0.5 M for DR antigen (J. Kaufman, personal communication) and between 0.1 M and 0.3 M for DC-1 antigen. Fractions containing the separated subunits were identified by running an aliquot on SDS-PAGE (12%) and pooled for sequence analysis.

I-A and I-E whose isoelectric points are quite different) and may share some antigenic determinants, they do not present any similarity in the region studied. More sequencing data is necessary to examine a possible common evolutionary origin. It is quite likely that these similar molecules have distinct functions.

We thank Deborah Shackelford, David Andrews and William Lane for advice and assistance. This work was supported by NIH grant AI-10736.

Note added in proof: The amino acid sequence of residues 17–232 of the DC α -chain has been obtained from a cDNA clone¹⁷ and, together with the data reported here, provides the complete sequence. This chain is highly homologous to the DR α -chain, with a notable exception at the N-terminus. Its α 2 region is a typical immunoglobulin-like domain. Moreover, amino acid sequence data have been presented which substantiate the conclusion that an I-A homologue is present on a DR 7 homozygous human cell line¹⁸.

Received 9 June; accepted 1 September 1982.

1. Strominger, J. L. *et al.* in *The Role of the Major Histocompatibility Complex in Immunobiology* (ed. Dorf, M. E.) 115–171 (Garland, New York, 1981).
2. McKusick, V. A. & Ruddle, F. H. *Science* **196**, 390 (1977).
3. Tosi, R., Tanigaski, N., Centis, N., Ferrara, G. B. & Pressman, D. *J. exp. Med.* **148**, 1592–1611 (1978).
4. Shackelford, D. A., Mann, D. L., van Rood, J. J., Ferrara, G. B. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4566–4570 (1981).
5. Shreffler, D. C., David, C. S., Cullen, S. E., Frelinger, J. A. & Neiderhuber, J. E. *Cold Spring Harb. Symp. quant. Biol.* **41**, 477 (1976).
6. McMillan, J., Cecka, J. M., Murphy, D. B., McDevitt, H. O. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5135–5139 (1977).
7. Allison, J. P. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 3953–3956 (1978).
8. Corte, G. *et al.* *Nature* **292**, 357–360 (1981).
9. Batchelor, J. R. & Morris, P. J. in *Histocompatibility Testing 1977*, 205–258 (Munksgaard, Copenhagen, 1978).
10. Terasaki, P. I. in *Histocompatibility Testing 1980*, 18–20 (UCLA Press, Los Angeles, 1980).
11. Duquesnoy, R., Marrari, M. & Vieira, J. in *Histocompatibility Testing 1980*, 861–863 (UCLA Press, Los Angeles, 1980).
12. Shackelford, D. A., Kaufman, J. F., Korman, A. J. & Strominger, J. L. *Immun. Rev.* **66**, 133–187 (1982).
13. Lampson, L. A. & Levy, R. J. *Immun.* **125**, 293–299 (1980).
14. Brodsky, F. M., Parham, P. & Bodmer, W. F. *Tissue Antigens* **16**, 30–48 (1980).
15. Silver, J., Cecka, J. M., McMillan, M. & Hood, L. *Cold Spring Harb. Symp. quant. Biol.* **41**, 369 (1977).
16. Goyert, S. M. & Silver, J. *Nature* **294**, 266–268 (1981).
17. Auffray, C., Korman, A. J., Roux-Dosseto, M., Bono, R. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **79** (in the press).
18. Goyert, S. M., Shively, J. E. & Silver, J. *J. exp. Med.* **156**, 550–566 (1982).
19. March, S. C., Parikh, I. & Cuatrecasas, P. *Analyt. Biochem.* **60**, 149–152 (1974).
20. Kowal, R. & Parsons, R. *Analyt. Biochem.* **102**, 72–76 (1980).

Differential expression of steroid sulphatase locus on active and inactive human X chromosome

Barbara R. Migeon*, Larry J. Shapiro†, Robert A. Norum‡, Thuluvancheri Mohandas†, Joyce Axelman* & Rebecca L. Dabora*

* Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

† Division of Medical Genetics, Harbor-UCLA Medical Center, Torrance, California 90509, USA

‡ Department of Medicine, Henry Ford Hospital, Detroit, Michigan 48202, USA

The X chromosome in mammalian somatic cells is subject to unique regulation—usually genes on a single X chromosome are expressed while those on other X chromosomes are inactivated¹. The X-locus for steroid sulphatase (STS; EC 3.1.6.2), the microsomal enzyme that catalyses the hydrolysis of various 3 β -hydroxysteroid sulphates, is exceptional because it seems to escape inactivation. Evidence for this comes from fibroblast clones in females heterozygous for mutations that result in a severe deficiency of this enzyme in affected males; all clones from these heterozygotes have STS activity, and enzyme-deficient clones that are expected if the locus were subject to inactivation², have not been found³. Further evidence that the STS locus escapes inactivation is that the human inactive X chromosome contributes STS activity to mouse-human hybrid cells⁴. On the basis of these hybrid studies^{5,6} the STS locus has been mapped to the distal half of the short arm (p22–pter) of the human X chromosome. Although the STS locus on both X chromosomes in human female cells is expressed, quantitative measurements of STS activity in males and females do not accurately reflect the sex differences in number of X chromosomes^{7–13} (Table 1). The ratio of mean values for normal females to that of normal males is greater than 1:1 but less than the ratio of 2:1 expected if STS loci on all X chromosomes were equally expressed. The incomplete dosage effect suggests that the STS locus on the inactive X chromosome might not be fully expressed. To test this hypothesis, we examined two heterozygotes for X-linked STS deficiency who were also heterozygous for the common electrophoretic variants of glucose-6-phosphate dehydrogenase (G6PD A and B). Studies of fibroblast clones from these females provide evidence, presented here, for differential expression of STS loci on the active and inactive X chromosome.

The affected males in the pedigree shown in Fig. 1 have X-linked ichthyosis¹³ and complete deficiency of microsomal STS activity, based on assays using cholesterol sulphate and dehydroepiandrosterone sulphate (DHEAS) as steroid substrates (Tables 2, 3). Heterozygotes II-4, II-7 and II-8 are obligate carriers of the mutant gene as their brother and sons are affected. Heterozygotes II-4 and II-7 are also heterozygous

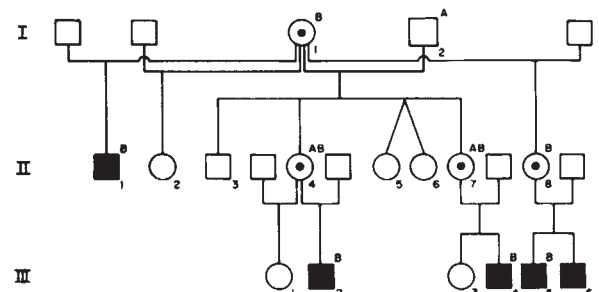


Fig. 1 Pedigree showing segregation of X-linked steroid sulphatase deficiency and G6PD variants. ■, Males with STS deficiency and X-linked ichthyosis. □, Normal male. ●, Carrier of STS deficiency. A, B indicate G6PD genotype.

Table 1 Comparison of specific activity of steroid sulphatase in male and female cells from reported studies

Tissue assayed	No. of specimens		Female: male activity ratio	Ref.
	Female	Male		
Fibroblasts	10	10	1.7	7
Fibroblasts	9	9	1.4	8
Fibroblasts	12	8	1.64	9
Placenta	18	18	1.8	10
Placenta	24	30	1.38	9
Lymphocytes	10	10	1.3	7
Leukocytes	25	12	1.3	11
Hair roots	9	8	1.04	12

Assays were carried out using ^3H -dehydroepiandrosterone sulphate.

for G6PD, having received the A allele from their father and the B allele from their mother. Therefore, in both females the STS mutation is on the same chromosome that carries the G6PD B allele.

Fibroblast cultures were established from skin biopsies of affected males and heterozygotes, and were maintained in Eagle's minimal essential medium supplemented with 15% fetal calf serum and non-essential amino acids. Control cultures included fibroblasts that had been similarly established from two females heterozygous for G6PD A, but normal with regard to the STS locus. Clonal cultures were established² from cells in their second subculture. Ten cells were plated into each of a series of 60-mm Petri dishes. Colonies were isolated 10 days later with cloning cylinders, transferred to establish clonal cultures and each clone was then assayed for G6PD phenotype by cellulose acetate gel electrophoresis¹⁴. Several clones of each G6PD type were transferred to flasks and delivered quickly to Torrance where steroid sulphatase was measured using ^3H -cholesterol sulphate¹⁵; determinations were carried out using two different amounts of cell extract in duplicate for each specimen to ensure linearity of the assay. For some specimens, aryl sulphatase A was assayed¹⁶ in the same lysate used for the STS and protein assays, while control assays of X-linked G6PD and the autosomal enzyme 6-phosphogluconate dehydrogenase (6PGD) were done¹⁴ on replicate cultures in Baltimore. Cultures were collected 10–11 days after plating and cell pellets were maintained at -20°C until assayed for sulphatases. The G6PD and 6PGD assays were carried out on fresh cell extracts. In each case, quantitative determinations were done without knowledge of origin or G6PD phenotype of the specimens.

In contrast to cells from the three affected males that had no detectable STS activity, all the clones from heterozygotes II-4 and II-7 showed STS activity (Table 2, Fig. 2), confirming observations in other heterozygotes³. However, the quantity of STS activity differed significantly in the two populations of clones from these double heterozygotes. The clones of G6PD B phenotype in each heterozygote had less activity than the G6PD A clones, and less activity than any of the clones from

the female controls. Although a few G6PD A clones in heterozygote II-4 had low STS activity, most were in the range of values for clones from the normal female, whereas all the G6PD B clones for II-4 had <4 pmol per mg protein per h of cholesterol sulphatase activity (Fig. 2). The results were similar when additional clones from heterozygote III-7 were assayed using ^3H -DHEAS as the steroid substrate (Table 3); the G6PD B clones had approximately half the activity of the G6PD A clones.

This differential activity was limited to the STS locus as activities for G6PD, 6PGD and aryl sulphatase A did not differ significantly between the two populations of clones in heterozygote II-4 (Table 2, Fig. 2). Because the activity of other enzymes is similar in both populations of clones, the differences in STS activity that we observed cannot be attributed to nonspecific differences in the metabolic activity of these cells. As the activity of STS is reduced in only one population of clones (those expressing G6PD B in both heterozygotes), it is unlikely that the decreased expression results from feedback inhibition of enzyme activity. The most probable explanation for our observations is that in both heterozygotes, the clones expressing G6PD B in fact produce less STS activity than clones that express G6PD A.

As there is no measurable enzyme activity for either steroid substrate in affected males (Tables 2, 3), the mutant locus in these individuals does not produce STS activity. Because the STS mutation in both heterozygotes is carried on the same chromosomes as the G6PD B allele, the active X chromosome in the G6PD B clones does not contribute STS activity. Therefore, all the STS activity in these clones must come from the normal allele on the inactive X chromosome. Furthermore, as G6PD B clones have less STS activity than G6PD A clones, it is clear that the normal STS allele produces less enzyme when it is on the inactive X, than on the active X chromosome.

Although our experimental observations are limited to heterozygous females from one family, we believe that they are representative of heterozygotes in other families also. The STS values for clones from heterozygotes reported previously (see Fig. 1 of ref. 3) are variable and their distribution is compatible with differential expression of the two X chromosomes in these females. Moreover, reduced expression of the STS locus on the inactive X chromosome is consistent with the incomplete dosage effect observed at this locus in normal individuals (see Table 1).

The reduced expression of the STS locus on the inactive X chromosome may reflect merely the influence of neighbouring genes that are subject to inactivation. Presumably, the gene is transcribed at a time when the inactive X chromosome has little if any other transcriptional activity, and is morphologically recognized as heterochromatin. Proximity to inactive X chromatin has been shown to inactivate neighbouring autosomal genes^{17–19}. On the other hand, differential expression at

Table 2 Enzyme activity in fibroblasts from normal females, affected males and heterozygotes for X-linked STS deficiency

Subject	Cloning efficiency	Specimen (no.)	Cholesterol sulphatase (pmol per mg per h)	Aryl sulphatase A (η mol per mg per h)	G6PD (μ mol per mg per min)	6PGD (μ mol per mg per min)
T. W., normal ♀	0.72	G6PD A clones (6)	15.3 $\pm$ 2.1	408 $\pm$ 75	0.175	0.038
		G6PD B clones (6)	15.5 $\pm$ 6.3	491 $\pm$ 146	0.187	0.036
	0.47	G6PD A clones (10)	11.0 $\pm$ 6.2	413 $\pm$ 137	0.169 $\pm$ 0.03	0.036 $\pm$ 0.00
		G6PD B clones (10)	2.4 $\pm$ 1.0	352 $\pm$ 76	0.201 $\pm$ 0.03	0.035 $\pm$ 0.01
W. K., normal ♀	0.36	G6PD A clones (6)	6.7 $\pm$ 1.0			
		G6PD B clones (6)	9.4 $\pm$ 2.2			
	0.39	G6PD A clones (8)	8.2 $\pm$ 2.9			
		G6PD B clones (10)	3.4 $\pm$ 0.6			
II-1, Affected ♂	0.39	G6PD B clones (8)	0.01 $\pm$ 0.0*			
II-1, Affected ♂		G6PD B, uncloned (3)	0.0*	320 $\pm$ 84		
III-4, Affected ♂		G6PD B, uncloned (2)	0.0	482		
III-5, Affected ♂		G6PD B, uncloned (2)	0.0	289		

Values of enzyme activity are mean $\pm$ s.d., or mean of two specimens if no s.d. is given. Cloning efficiency, clones obtained/cells plated. Brackets indicate specimens assayed simultaneously.

* Note that the absence of activity in affected males indicates that cholesterol sulphate is a specific substrate for STS.

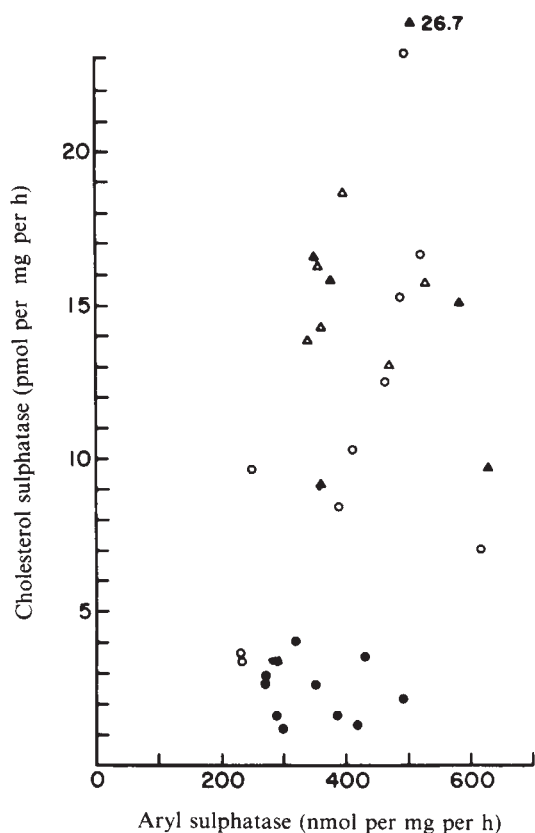


Fig. 2 STS values for individual clones as a function of aryl sulphatase A activity. ○, G6PD A and ●, G6PD B clones from heterozygote II-4; △, G6PD A and ▲, G6PD B clones from T.W., the normal female control.

the *STS* locus could be due to other mechanisms regulating the expression of genes on the X chromosome. Having a single active X chromosome in mammalian cells of both sexes serves to equalize the expression of X-linked genes in male and female somatic cells, but results in an effective monosomy for these genes. The method of X chromosome dosage compensation in *Drosophila* avoids this problem because sexual equality is achieved by increasing the transcriptional activity of the single X in the male²⁰ rather than by inactivating one of the two X chromosomes in the female. Thus, males as well as females have effectively two doses of each X gene. Like the dipteran chromosome, the mammalian X chromosome carries many genes having functions similar to those on autosomes. As monosomy is usually detrimental for mammals, it is possible that mechanisms have evolved to compensate for the single copy of X genes, by enhancing the product output of the active X chromosome²¹⁻²³. The greater activity of the *STS* locus on the active X chromosome observed here could be attributable to gene duplication²² (inconsistent with genetic evidence) or to increased output of the active X locus through enhanced transcription or processing.

Further evidence for greater activity of active X loci comes from studies of normal human fibroblasts, heterozygous for

G6PD A, that have undergone spontaneous derepression of the G6PD locus on the inactive X chromosome²⁴. Electrophoretic analysis of these cells indicates that the reactivated locus on the inactive X contributes half as many subunits as that on the active one. Therefore, it seems that the output of X loci that escape inactivation early in normal development or that are expressed following derepression, is significantly less than that of corresponding loci on the active X. Although compatible with some mechanism that down-regulates expression of the inactive locus, the differential activity may reflect increased expression of the active one. In either case, our observations indicate additional complexity in the regulation of X-chromosomal loci in man.

This work was supported by NIH grants HD 05465 (B.R.M.) and HD 12178 (L.J.S.) and March of Dimes Birth Defects Foundation grant 1-639 (L.J.S.).

Received 28 May; accepted 2 September 1982.

1. Lyon, M. F. *Biol. Rev.* **47**, 1-35 (1972).
2. Migeon, B. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 5066-5070 (1981).
3. Shapiro, L. J., Mohandas, T., Weiss, R. & Romeo, G. *Science* **204**, 1224-1226 (1979).
4. Mohandas, T., Sparkes, R. S., Hellkuhl, B., Grzeschik, K. H. & Shapiro, L. J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6759-6763 (1980).
5. Mohandas, T., Shapiro, L. J., Sparkes, R. S. & Sparkes, M. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5779-5783 (1979).
6. Müller, C. R., Westerveld, A., Migl, B., Franke, W. & Ropers, H. H. *Hum. Genet.* **54**, 201-204 (1980).
7. Müller, C. R., Migl, B., Traupe, H. & Ropers, H. H. *Hum. Genet.* **54**, 197-199 (1980).
8. Ropers, H. H. *et al. Hum. Genet.* **57**, 354-356 (1981).
9. Bedin, M., Weil, D., Fournier, T., Cedard, L. & Frezal, J. *Hum. Genet.* **59**, 256-258 (1981).
10. Lykkesfeldt, G., Bock, J. E. & Lykkesfeldt, A. E. *Lancet* **ii**, 255-256 (1981).
11. Epstein, E. H. Jr & Leventhal, M. E. *J. clin. Invest.* **67**, 1257-1262 (1981).
12. Dancis, J., Jansen, V. & Hutzler, J. *Pediat. Res.* **14**, 521 (1980).
13. Shapiro, L. J., Weiss, R., Webster, D. & France, J. T. *Lancet* **i**, 70-72 (1978).
14. Maren, C. & Migeon, B. R. *Am. J. hum. Genet.* **33**, 752-761 (1981).
15. Shapiro, L. J. *et al. Lancet* **ii**, 756-757 (1978).
16. Baum, H., Dodgson, K. S. & Spencer, B. *Clinica chim. Acta* **4**, 453-455 (1959).
17. Cattana, B. M. Z. *VererbLehre* **92**, 165-182 (1961).
18. Eicher, E. *Adv. Genet.* **15**, 175-259 (1970).
19. Mohandas, T., Sparkes, R. S. & Shapiro, L. J. *Am. J. hum. Genet.* (in the press).
20. Belote, J. M. & Lucchesi, J. C. *Nature* **285**, 573-575 (1980).
21. Ohno, S. *Sex Chromosomes and Sex-linked Genes*, 99 (Springer, New York, 1967).
22. Lyon, M. F. *Proc. R. Soc. B* **187**, 243-268 (1974).
23. Lucchesi, J. C. *Science* **202**, 711-716 (1978).
24. Migeon, B. R., Wolf, S. F., Maren, C. & Axelman, J. *Cell* **29**, 595-600 (1982).
25. Hameister, H. *et al. Hum. Genet.* **46**, 199-207 (1979).

Structure of the S-layer of *Sulfolobus acidocaldarius*

K. A. Taylor*, J. F. Deatherage & L. A. Amos

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The outermost component of many bacterial cell envelopes is a two-dimensional crystalline array of protein molecules, termed the S-layer¹. Despite considerable effort to investigate its structure and composition, its precise function remains uncertain. We report here the first determination of the three-dimensional structure of an S-layer, that from the thermophilic bacterium, *Sulfolobus acidocaldarius*. Our map shows a complex and strongly interconnected structure, penetrated by numerous channels which appear to provide little barrier to anything but rather large molecules. The external surface is fairly smooth, but the cellular side is sculpted with large cavities and protruding 'pedestals'. The protein substructure consists of three types of globular domain, connected by narrow bridges. We suggest that these may be flexible 'hinges', allowing the S-layer to form a curved surface, and that the multidomain structure allows the S-layer to maintain strong connectivity as it grows with the bacterium.

S. acidocaldarius is a facultative sulphur-oxidizing autotroph whose natural habitat is hot acidic springs². Cells are roughly spherical, 0.8-1.0 µm in diameter, and are often lobed. When

Table 3 Steroid sulphatase activity in fibroblasts assayed with DHEAS as substrate

Subject	Specimen (no.)	DHEAS sulphatase (pmol per mg per h)
II-7, Heterozygote	G6PD A clones (4)	2,665 ± 433
	G6PD B clones (4)	1,518 ± 432
III-2, Affected ♂	Uncloned	0
	Uncloned	3,650

Cells were washed, pelleted and suspended in buffer with sodium lauryl sulphate¹⁵ and disrupted by sonication (15 s). The reaction was carried out according to Hameister²⁵ in duplicate at two substrate concentrations using ³H-DHEAS (22.1 Ci mmol⁻¹; NEN) in a final concentration of 6.3 × 10⁻⁶ M at 37 °C for 2 h. The desulphated substrate was recovered after two benzene extractions¹¹.

* Present address: Department of Anatomy, Duke University Medical Center, Durham, North Carolina 27710, USA.

grown in autotrophic conditions with elemental sulphur as an energy source, the cells attach to sulphur crystals by means of pili³. Like other members of the Archaeobacteria, they have a simple cell wall, consisting essentially of the S-layer attached to the plasma membrane⁴. The S-layer consists of a single glycoprotein with an apparent molecular weight (MW) of 140,000–170,000, as determined by SDS-gel electrophoresis⁵. In the electron microscope, untilted specimens of purified S-layer show a hexagonal lattice, two-sided plane group p6, with a 220 Å unit cell dimension (Fig. 1a).

The three-dimensional model (Fig. 2) is more complex than would be expected from direct inspection of the two-dimensional projection of the structure (Fig. 1b), which apparently shows fairly simple dimer units arranged to form a series of hexagonal and triangular holes. In fact, although the three-dimensional structure is very open, with protein occupying just 30% of the volume within the ~100 Å-thick layer, only the holes on the 6-fold axes pass directly through. One side of the layer is relatively smooth, the other deeply indented. Shadowing of intact cells⁶ and isolated S-layers⁷ has shown that the smooth side is the exterior surface. The structure consists essentially of three types of oligomeric globular domains which we have named the diad (D), triad (T) and 6-fold ring (R) regions (Fig. 2). Six large D domains around each ring form a domed chamber on the cellular side of each 6-fold axis (Fig. 2b). At the top of the dome is a hole of 50 Å diameter. On the 3-fold axes, around the rims of the chambers, are the pedestal-like T domains, which presumably bind the S-layer to the cell surface. Published micrographs⁶ suggest that these features are in contact with the surface of the membrane, rather than being in the lipid bilayer.

At this resolution, it is impossible to determine molecular boundaries. At least two monomers must contribute to the D domain, at least three to the T domain and six to the ring. The simplest possible molecular shape (illustrated in Fig. 2) is nevertheless complicated in terms of intermolecular connections, making a total of eight separate contacts with six different neighbours. Indeed, the division into separate polypeptides is likely to be more complex and the number of intermolecular interactions even greater.

Whatever the shape of the individual monomer, it is useful to consider the assembled structure in terms of the morphological features and their interactions. Connecting the D and T domains is a relatively narrow waist (H1) which we believe could function as a flexible hinge. Similarly, two different types of narrow waists (H2, H3) connect the D and R domains. Until the shape of the S-layer monomer is known, it is impossible to determine whether either of the latter connections is a non-covalent intermolecular interface as opposed to a covalent link or hinge. The outer surface is made up of R domains, which terminate in numerous small knobs. Adjacent R domains are linked by a fourth kind of narrow connecting region (H4), which is long and has two sharp bends. Flexing of these bends would allow the distance between adjacent rings to vary, concertina-fashion, and accommodate the distortions necessary for lobe formation (see below).

An analogous multidomain structure connected by hinges has been found for the coat protein subunit of tomato bushy stunt virus (TBSV), which consists of two domains (P and S) connected by a hinge⁷. The TBSV subunit occupies three quasi-equivalent positions in the virus shell, yet the polypeptide chain foldings within the three different kinds of P and S domains are almost identical, the major differences between quasi-equivalent subunits being the angle between the domains. This seems likely to be a common feature of allosteric structural proteins, although as yet there is little evidence as to few structural proteins have been studied at sufficiently high resolution. An S-layer built on this principle could combine strength and flexibility having rigid domains connected by covalently-bonded hinges. Molecules could thus assume a range of similar but not identical conformations as the S-layer curves around the bacterial surface.

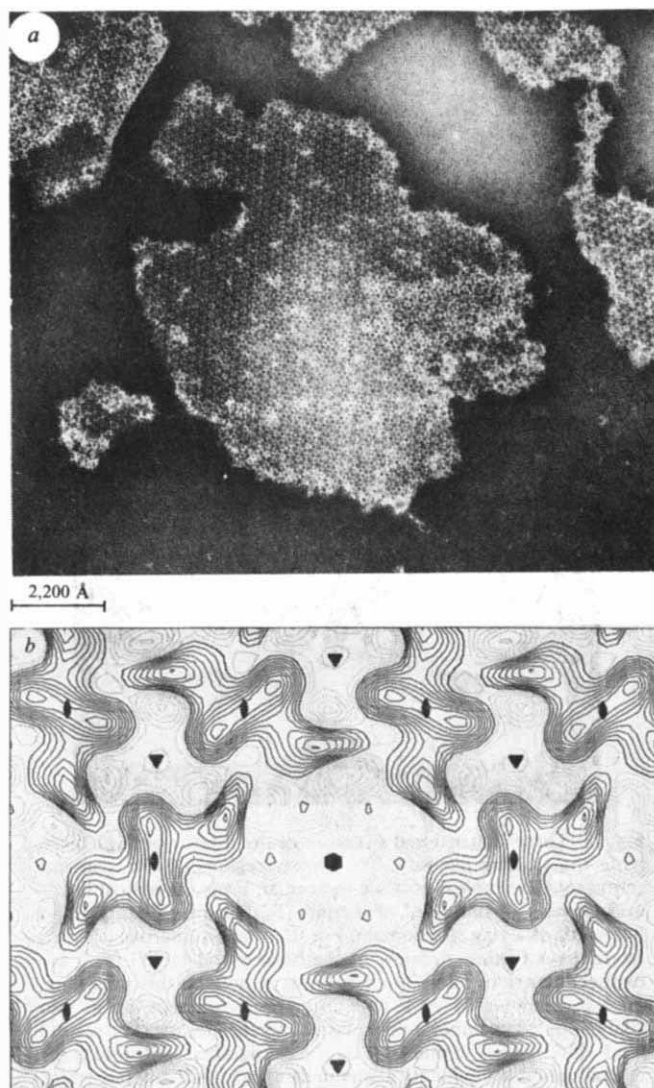


Fig. 1 a, Cell wall sacculi were isolated from *S. acidocaldarius*, strain 98-3, by procedures described elsewhere⁵. These were broken into small flat sheets by brief sonication. Specimens for electron microscopy were prepared on glow-discharged carbon support films using potassium phosphotungstate, pH 7.0 as negative stain and observed in a Philips EM 400 electron microscope equipped with goniometer stage and low dose kit for minimizing electron exposure. A specimen in fairly thick stain is shown here for good contrast (most of the images processed were more lightly stained). The relative sizes of the apparent holes in 3-fold and 6-fold positions varies according to the depth of stain. b, In-plane projection of a small part of the *Sulfolobus* S-layer. Dark contours represent protein, light contours negative stain. There are apparently large holes in the 6-fold positions and smaller holes in 3-fold positions. The unit cell size is 220 Å.

On a curved surface, the subunits must either move more closely together on the inside or farther apart on the outside, compared with the flat structure we have solved. The former possibility is the more likely, as having the T-domain pedestals spread well apart on the membrane side would prevent them from interfering with one another if the spacing were reduced. At junctions between lobes, the wall would curve the other way, with the T domains spreading farther apart than in a flat structure.

However, a single crystalline S-layer cannot form a closed surface simply by bending. Closure could be achieved in crystalline shells by the introduction of edge dislocations⁸, examples of which have been shown in various other species^{9,10}. In the case of layers having hexagonal symmetry, however, Caspar and Klug¹¹ have discussed the possibility that closed containers may be formed by introducing pentagons instead of hexagons,

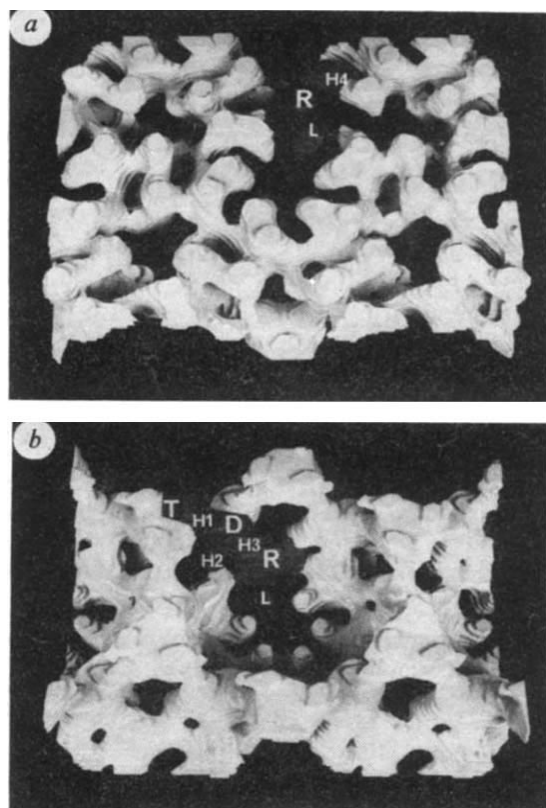


Fig. 2 Three-dimensional structure of the S-layer of *Sulfolobus*. Note the striking porosity. The darker regions of the structure represent the simplest possible molecular shape, an almost linear arrangement of one-third of a triad (T), half of a diad (D) and one-sixth of a ring (R) domain, but there are numerous possible alternatives to this. *a*, View of the balsa wood model from the extracellular side. There are two kinds of hole on this side. One, 50 Å in diameter, in the R domain on the 6-fold axis, opens into a large domed chamber on the plasma membrane side. The second kind of hole, 45 Å in diameter on the 3-fold axes, splits into three crystallographically identical channels of ~25 Å diameter, which open into the sides of the large chamber. *b*, View from the plasma membrane side. The chamber on the 6-fold axis is ~185 Å wide, measured between T domains, and ~80 Å deep. The three-dimensional structure was determined using procedures similar to those of Henderson and Unwin¹⁹, except that both phases and amplitudes were obtained from the computed Fourier transforms of the images. Data from three 10-member tilt series, to angles as high as 60°, and 12 fixed-angle views between 72° and 78° were combined. The resulting three-dimensional model contains almost all the data to 17 Å resolution with the exception of the F(00L) and part of the F(10L). The average phase residual for the three-dimensional origin refinement was 15–25.3 degrees. Variation in density across flat negatively stained areas is 6–10% of the variation across stained protein boundaries. Further details of the data processing will be published elsewhere²⁰. The cut-out model has 95% of the volume predicted for a monomer of MW 170,000.

to give a quasi-equivalent arrangement analogous to that obtained with icosahedral viruses. Edge dislocations are then unnecessary and a stronger wall with no unsatisfied contacts may be achieved without them. To close the spherical shell in this manner requires a minimum of 12 pentagons. By extending this principle we believe that the phenomenon of lobe formation in *S. acidocaldarius* may be explained through the insertion of additional pentagons and compensating heptagons into the surface lattice. Note that surface lattices involving mixtures of pentagons, hexagons and heptagons have been demonstrated, indeed they have been observed at various stages in coated pit formation¹².

The basket-like network raised away from the membrane on pedestals is one of the most striking features of the reconstruction. Several lines of evidence suggest that this may be a fairly

general property of S-layers: for example, the model proposed for the S-layer of *Spirillum serpens*¹³ has similar features, although it is formed in an entirely different fashion. Moreover, edge-on views of negatively stained S-layers, as observed for *Acinetobacter*¹⁴ and profiles viewed in thin section from such species as *Pelodictyon clathratiforme*¹⁵ and *Methylobacter*¹⁶ also suggest a similar feature in their three-dimensional structure.

Several physiological functions have been proposed for bacterial S-layers¹. Protection and support of the cell are obvious possibilities. This seems particularly likely in the case of the Archaeobacteria, in which the S-layer seems to be the only component of the cell wall. In *S. acidocaldarius*, it may be essential for providing support against the low osmotic pressure in hot acid springs. In this case, the strength of the S-layer would need to be maintained during cell growth, even while additional molecules are being inserted into the layer. The need to provide a constant framework of support, while allowing a degree of flexibility and porosity, may explain why the subunits are so complex and make so many separate intermolecular contacts. Extra rows of molecules might be inserted by a multi-step process, whereby different sets of bonds would be broken and transferred to the new subunits in separate stages, so that at no time would there be a complete break in the S-layer. By an elaboration of this insertion process, involving the introduction of additional pentagons and compensating heptagons, the S-layer could adapt to lobe formation or to cell division without weakening.

A possible function attributed to the S-layer is that of a selective permeability barrier. Stewart and co-workers^{17,18} have recently suggested, on the basis of projected views of the S-layers from *Aquaspirillum putridiconchylum* and *Sporosarcina ureae*, that they might function as molecular sieves, allowing the passage of small essential metabolites but not proteins. However, the channels we observe in our three-dimensional model of the *Sulfolobus* S-layer seem too large to be selective in this way. This S-layer may be designed simply to provide the cell with as much access to the medium as possible, while still maintaining its supportive function. It may also be relevant that the diameter of the channel through the S-layer on the 6-fold axis is very similar to the size of the pili which attach the bacterium to sulphur crystals³.

We have found the S-layer of *S. acidocaldarius* to be an excellent specimen for three-dimensional image reconstruction. We hope, therefore, that similar studies on other examples will provide further clues to the mode of assembly and function of cell walls.

We thank H. Michael for providing *S. acidocaldarius* cell sacculi and R. Henderson and all those who participated in the EMBO course on 'Three-Dimensional Structure Analysis by Electron Microscopy', held in Cambridge in 1980, for their help and encouragement. K.A.T. was supported by NIH grant R01 GM/AM 28224 and was assisted by a Research Travel Grant from the Burroughs-Wellcome Fund.

Received 24 May; accepted 20 August 1982.

1. Sleytr, U. B. *Int. Rev. Cytol.* **53**, 1–64 (1978).
2. Brock, T. D. *Thermophilic Microorganisms and Life at High Temperatures*, 117–179 (Springer, New York, 1978).
3. Weiss, R. L. *J. gen. Microbiol.* **77**, 501–507 (1973).
4. Weiss, R. L. *J. Bact.* **118**, 275–284 (1974).
5. Michel, H., Neugebauer, D.-Ch. & Oesterheld, D. in *Electron Microscopy at Molecular Dimensions* (eds Baumeister, W. & Vogell, W.) 27–35 (Springer, New York, 1980).
6. Brock, T. D., Brock, K. M., Bely, R. T. & Weiss, R. L. *Arch. Mikrobiol.* **84**, 54–68 (1972).
7. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* **276**, 368–373 (1978).
8. Harris, W. F. & Scriven, L. E. *Nature* **228**, 827–829 (1970).
9. Watson, S. W. & Remsen, C. C. *Science* **163**, 685 (1969).
10. Sleytr, U. B. & Glaeser, A. M. *J. ultrastruct. Res.* **50**, 103–116 (1975).
11. Caspar, D. L. D. & Klug, A. *Cold Spring Harb. Symp. quant. Biol.* **27**, 1–24 (1962).
12. Heuser, J. J. *Cell Biol.* **84**, 560–583 (1980).
13. Glaeser, R. M., Chiu, W. & Grano, D. *J. ultrastruct. Res.* **66**, 235–242 (1979).
14. Thorne, K. J. I. *Biol. Rev.* **52**, 219–234 (1977).
15. Cohen-Bazire, G., Kunisawa, R. & Pfennig, N. *J. Bact.* **100**, 1049–1061 (1969).
16. Wilkinson, J. F. *Symp. Soc. gen. Microbiol.* **21**, 15–46 (1971).
17. Stewart, M. & Beveridge, T. J. *J. Bact.* **142**, 302–309 (1980).
18. Stewart, M., Beveridge, T. J. & Murray, R. G. E. *J. molec. Biol.* **137**, 1–8 (1980).
19. Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28–32 (1975).
20. Deatherage, J. F., Taylor, K. A. & Amos, L. A. *J. molec. Biol.* (submitted).

BOOK REVIEWS

Lemurs, dead and alive

R.D. Martin

THE prosimians, commonly referred to as "lower primates", can be loosely regarded as representing a primitive grade closer to the ancestral primate condition than the modern monkeys and apes. For this reason, any study of primate evolutionary biology is particularly dependent upon an understanding of the prosimians (lemurs; lorises; tarsiers). Yet, until quite recently, the prosimians were relatively neglected in the primate literature, though several recent publications have done much to fill this gap.

Among the prosimians, the Madagascar lemurs comprise by far the largest assemblage of living species and provide a remarkable example of adaptive radiation on an isolated land-mass. Until now, however, there has been no really comprehensive treatise on the Madagascar lemurs available in the English language; most of the relevant publications are in French. Now, at long last, a highly readable and authoritative English-language reference work on the lemurs has been provided in the form of Ian Tattersall's *The Primates of Madagascar*. This (by today's standards) reasonably priced and attractively presented book is a must for libraries and research workers concerned with primate evolutionary biology.

After providing a neat and illuminating historical survey of the discovery and early description of Madagascar's lemurs, Tattersall goes on to consider the general biogeographical background, giving due reference to continental drift. It now seems to be well established that Africa was the primary source of the original mammal fauna which colonized Madagascar and that this large oceanic island has existed as such for at least the past 65 million years. Adaptive radiation of the lemurs to yield the modern array of a score or so species undoubtedly took place largely or exclusively within Madagascar. These living lemur species are effectively reviewed by Tattersall, who provides standard measurements from his comprehensive survey of major museum collections and body weights wherever possible, as well as providing useful distribution maps incorporating actual locality records in addition to the customary expanses of hopeful shading.

It is a sad fact that until historical times Madagascar's lemur fauna was almost twice as rich in species as it is now, including numerous forms which were markedly bigger than most extant lemurs,

The Primates of Madagascar. By Ian Tattersall. Pp.382. ISBN 0-231-04704-5. (Columbia University Press: 1982.) \$52.

such as the ape-sized *Megaladapis* and the monkey-like *Archaeolemur*. These large-bodied lemurs are known only as subfossils and Tattersall has performed a particularly valuable task in providing a chapter with the clearest available general review of the subfossil lemurs. He also gives a balanced discussion of conflicting explanations for the recent extinction of these animals (human intervention versus climatic change).

The remaining chapters provide very useful syntheses of information on behaviour and ecology, general morphology and adaptation, and phylogeny and classification. The thoughtful concluding chapter ends, as it should, with a strong statement about increasing threats to the survival of the remaining lemur species.

Tattersall's central research interest has been the phylogenetic relationships of the lemurs, and he provides a fairly comprehensive survey of relevant information and conflicting interpretations. His conclusions, however, are somewhat controversial. In agreement with several other authors, he cites morphological evidence (essentially from the auditory bulla and internal carotid circulation) suggesting that certain lemurs (family Cheirogaleidae) diverged from a common ancestry with the Afro-Asian loris group (family Lorisidae) post-dating the initial radiation of the lemurs. Serological evidence, on the other hand, supports the alternative view that the lemurs of Madagascar constitute a self-contained (monophyletic) group.

Tattersall dismisses the biochemical and chromosomal evidence without proper discussion, and this — along with a general reluctance to emphasize quantification — is one of the few shortcomings of an otherwise well-balanced book. It is, of course, impossible to cover everything in a single text, but the principle must be recognized that any overall assessment of lemur phylogeny must eventually integrate biochemical and chromosomal data. The hypothesis of a later link between the Cheirogaleidae and the Lorisidae automatically requires at least two crossings of the Mozambique Channel by early lemurs. Tattersall goes still further than this in postulating, on craniodental grounds, specific relationships between Eocene fossil adapids and certain extant lemurs

(*Notharctus* + *Lepilemur*; *Adapis* + *Hapalemur*). This would require not only multiple crossings of the Mozambique Channel but also a very early origin for the Madagascar lemur radiation. The dental similarities between the Eocene fossils and modern lemurs are, indeed, striking; but parallel evolution would seem to be a strong possibility. Nonetheless, Tattersall's views on lemur phylogeny are thought-provoking and will doubtless fuel further research in this area.

Overall, *The Primates of Madagascar* performs an excellent service in providing an up-to-date review of a wide range of material on the Madagascar lemurs, together with a comprehensive bibliography. It will undoubtedly become a standard reference work for future studies of these unique primates. □

R.D. Martin is Professor of Physical Anthropology at University College London.

Grains of evidence

A.T. Grove

Environmental History of East Africa: A Study of the Quaternary. By A.C. Hamilton. Pp.328. ISBN 0-12-321880-2. (Academic: 1982.) £24, \$44.50.

STUDIES of Late Quaternary environmental history in East Africa exploit the opportunities afforded by the sedimentary record and old shorelines in the rift basins, and by glaciation features on the high mountains. Pollen preserved in the sediments of the rift and in cirque basins, and also in peats in the higher and wetter areas, is an indicator of former plant distributions. It is with the last subject that Hamilton is mainly concerned in this book. Extraction, identification and counting of the pollen grains is testing and time-consuming; interpretation requires a good general knowledge of plant geography. It remains, I suspect, an art as well as a science.

Hamilton has reduced some of the uncertainties of interpretation in East Africa by examining surface pollen samples in relation to current vegetation patterns and drawing conclusions as to the production and mobility of different types of pollen. This has enabled him to

Essential information at your fingertips Every month!

The British Journal of Clinical Pharmacology is the leading authoritative journal presenting results of clinical pharmacological research.

Published monthly on behalf of the British Pharmacological Society the journal:

- ★ Bridges the gap between medical professions, clinical research and the Pharmaceutical industry.

- ★ Contains papers and reports on all aspects of drug action in man, in both health and disease.

- ★ Prints letters and short communications which allow immediate discussion of urgent observations.

Subscribers to British Journal of Clinical Pharmacology receive, free of charge, the proceedings of selected symposia on new methods, new drugs, and new approaches to treatment. Some past symposia have been on narcotic antagonists and antagonist analgesics, labetalol, clobazam, mianserin, temazepam and the relationship between angina pectoris and hypertension.

Why not join the growing number of clinical pharmacologists, clinicians and others who rely on the British Journal of Clinical Pharmacology as their information medium.

For a sample copy write to: Felicity Parker, Scientific & Medical Division, Macmillan Publishers Ltd., Houndmills, Basingstoke, Hampshire, RG21 2XS.

Circle No.63 on Reader Service Card.

reinterpret early pollen diagrams as well as to interpret his own more effectively. The results confirm the conclusions reached by other methods as to the Late Quaternary climatic succession in the region — a cool, dry Late Pleistocene period from about 20,000 to 12,000 BP being preceded and followed by wetter millennia, and current climates becoming established about 4,000 BP.

The significance of Late Quaternary climatic change for the history of the equatorial rain forest has been underlined by J.R. Flenley in *The Equatorial Rain Forest: A Geological History* (Butterworths, 1979). Flenley's field experience was mainly in south-east Asia, however, and evidence of such effects is greater in Africa. Hamilton has worked, on and off, for the last 16 years in East Africa and knows the highland areas intimately from his own field studies made, in recent years, in collaboration with R.A. Perrott.

In this book he considers the altitude zonation of the vegetation and presents, with interpretations, pollen diagrams from Elgon, Cherangani, Mount Kenya, Kilimanjaro, Badda and Bale in Ethiopia, Ruwenzori and the Rukiga Highlands in Uganda, the Kamiranzovu swamp in Rwanda and the dambos of the Nyika Plateau in Malawi (the last by M.E. Meadows). A feature of many of the cores is the absence of peat for some millennia before 4,000 BP; Hamilton explains this in terms of greater surface wash under the wetter conditions of the Early Holocene; but could it rather be the result of wastage accompanying desiccation about 4,000 to 4,500 BP?

Hamilton argues that the key to an understanding of the evolutionary processes in East African vegetation is the recognition of core areas of limited extent where rain forest survived the dry periods of the Quaternary. From these core areas in west Africa, Cameroun-Gabon, eastern Zaire and the Usumbara and Uluguru mountains near the east African coast the diversity of species diminishes outwards. He briefly reviews the past and probable future effects of man on the east African environment and recommends that conservation of the core areas should receive special attention.

This is a welcome addition to the literature on environmental change in low latitudes by one of the most active and experienced researchers in the field. Hamilton presents the botanical evidence bearing on the Quaternary history of East Africa clearly and simply enough to allow the non-specialist to appreciate the implications of the pollen records and see them within a more general setting. Earth scientists and ecologists at both the undergraduate and research level will find his study useful and stimulating. □

A.T. Grove is a Geographer at Downing College and Director of the African Studies Centre, University of Cambridge.

Old space dust

David W. Hughes

Interplanetary Dust. By P.W. Hodge. Pp.280. ISBN 0-677-03620-5. (Gordon & Breach: 1982.) \$38.

DUST gets everywhere and this is certainly true of the interplanetary variety. As it orbits the Sun, it scatters radiation, resulting in a cone of light extending along the ecliptic known as zodiacal light. It collides with the Earth, is brought to a halt in the atmosphere and drifts down to the surface where it can be recovered from deep-sea sediments, in Antarctic and Arctic ice, and as a component of airborne dust. It can be sampled using balloons, rockets and high-flying aircraft. Interplanetary dust continuously bombards the lunar surface producing microcraters in exposed glasses and rocks. It can be measured using satellite-borne instruments and was once regarded as a serious hazard to space missions. In the interplanetary context any particle smaller than 100 μm , about a millionth of a gram, is regarded as dust.

The history of dust investigation is fascinating. It is a science that got off to a bad start. Early satellite experiments both in the United States and the USSR agreed well — and fitted rocket data — but all produced results which were a factor of 10⁶ wrong. Microphones were detecting the creaking of spacecraft and not the impact of dust. Contamination has always been a problem; in airborne dust less than one particle in 10⁹ is extraterrestrial.

P.W. Hodge, the author of this book, played an important part in many of the early dust studies and has produced a short (about 45,000 word) summary of the subject — published in a camera-ready format. His book is easy to read, and provides a reasonable overview of the early days of the subject. I stress "early days" because the references tail off sharply after 1976. The standard of coverage is patchy and the concentrations of technicalities occur in the areas of micrometeorites, fossil dust collection and particles around meteorite craters — all subjects on which Hodge has worked. The illustrations are of a rather mixed quality and many of the photographs of the particles are printed without scales.

This is the sort of book I might recommend a research student to read if he had nothing to do for an afternoon. But I would then insist that he spent much, much longer with, for example, the IAU Symposium 90 report (*Solid Particles in the Solar System*, edited by I. Halliday and B.A. McIntosh and published by Reidel in 1980) to give him an up-to-date overview of the subject, a feel for the excitement of the work of the cosmic dust men and a taste for the potential avenues of research. □

David W. Hughes is a Lecturer in Astronomy and Physics at The University of Sheffield.

Endocrinology through the changes

J.G. Phillips

Hormones in Development and Aging. Edited by Antonia Vernadakis and Paola S. Timiras. Pp.686. ISBN 0-85200-560-1. (MTP Press: 1982.) £50.25.

REVIEW books on important and topical areas of biomedical research are always to be welcomed. Such a reception should greet *Hormones in Development and Aging*. Vernadakis and Timiras, the editors, have assembled an impressive group of authors to review the literature, highlight current concepts and anticipate developments in an area of endocrinology which holds great current interest both in its fundamental aspects and in its impact on clinical practice.

The overall approach has been to provide detailed coverage of the events running from fertilization through to differentiation, development, maturation and ageing, as well as of the role hormones play in organizing cellular expression and regulating metabolic events. Rightly central in the thinking of the authors has been the underscoring of the interrelationship of the endocrine and nervous systems, as both are crucially involved in the processes of growth and adaptation, puberty, rhythms of biological importance, the onset and ageing of reproductive functions, general ageing and neoplasia.

The secondary aim has been to consider the effects of hormones on brain development, both normal and abnormal, and to examine their influence on behaviour throughout the life span. This longitudinal approach allows each of the contributors to describe the basic aspects of changes in endocrinology during development and ageing, as well as pathological alterations which subtract from the optimum conditions for good health and the capacity to maintain vigour into "old age".

The authors have also seized upon the opportunity to propose interventive approaches wherever possible. This emphasis is in tune with the widely held view that, by finding means of strengthening physiological function, there will be a swing away from the current clinical practice of finding the cure for the disease towards the elaboration of a set of well understood guidelines which, when followed, will possibly serve to maintain adaptability and vigour throughout the whole life span. The clinical potential of such procedures is self-evident.

The impact of the book is unfortunately diminished by numerous errors in the text, a lack of consistency in presentation of data, poor photographs and an amazing paucity of very recent papers — indeed, it is explicitly stated that Ramaley's chapter, "Neuroendocrinology of Puberty", was written five years ago! On balance, however, this is a valuable compilation of

reviews which must find its way into libraries and onto the desks of gerontologists, geriatricians and developmental biologists. □

J.G. Phillips is Director of the Wolfson Institute at the University of Hull.

Before Einstein

Stephen G. Brush

Energy, Force, and Matter: The Conceptual Development of Nineteenth-Century Physics. By P.M. Harman. Pp.182. Hbk ISBN 0-521-24600-8; pbk ISBN 0-521-28812-6. (Cambridge University Press: 1982.) Hbk £13.20, \$27.50; pbk £5.50, \$8.95.

IT IS NO easy task to give a concise, accurate and readable account of nineteenth-century physics. Yet someone must undertake it if students and general readers are to gain an understanding of the background for the spectacular theories and discoveries of the twentieth century. In his contribution to the distinguished *Cambridge History of Science Series*, P.M. Harman (formerly Heimann) has satisfied the first two criteria admirably, the third passably. He has condensed an amazing amount of material into fewer than 200 pages (including a valuable bibliographic essay and a useful index); his facts and interpretations are in accord with the best modern scholarship; but his style is not lively enough to attract a reader who is not already convinced of the importance of the subject. It is probably the best textbook currently available for the first part of a course on the history of modern physics, as long as the instructor can motivate students to struggle through Harman's densely-written prose.

Following a rapid survey of imponderable fluids, the ether concept, Fourier's heat theory and the discovery of energy-conversion processes in the first half of the century, Harman gives a thorough treatment of the development of thermodynamics and field theories of electricity and magnetism. An account of Hertz's discovery of electromagnetic waves, the Michelson-Morley experiment and Lorentz's electrodynamics prepares the reader for the introduction of relativity; but, as a professional historian striving to avoid the Whig interpretation (judging the past only as progress towards the present), Harman carefully refrains from mentioning Einstein's theory at this point. Instead, he jumps back to Dalton's chemical atomic theory and traces

molecular physics through Clausius, Maxwell and Boltzmann. A final chapter ties together the late nineteenth-century critiques of atomism and the difficulties in ether theory accompanying the decline of the mechanical world view. Having pointed out the continuity between these trends of the 1890s and the revolutionary ideas of Planck and Einstein, the author ends his story just before the conflicts in his plot are resolved.

Readers familiar with current writings in the history of science will perhaps regret that Harman has chosen to minimize the philosophical, social and psychological aspects of physics; we find here no mention of Kuhnian paradigms, Lakatosian research programmes, or (a more recent fad) Mary Douglas's grid-group categories. This is probably just as well, however, since the instructors or the more sophisticated readers will probably prefer their own version of such schemes to whatever the author might present, and it is always useful to have one straightforward account of the basic technical developments.

For me, a more serious omission is the mathematical basis of physical theories. Incredible as it might seem to a physicist, Harman manages to discuss all the above-mentioned topics without using a single equation or mathematical symbol. While this policy may succeed in making the book appear accessible to a larger audience, I doubt if a phrase such as "Lagrangian formalism" will mean much to the reader who lacks knowledge of theoretical physics. On the other hand there is a definite advantage in having a mathematics-free text if the missing equations are presented and explained in a separate chapter at the end, so that those readers who wish to know, for example, what Maxwell's field equations are and how they explain the propagation of light, can find out.

For those authors who have been frustrated by the refusal of publishers to print discursive footnotes (as opposed to endnotes), the most exciting aspect of this book may be Harman's innovative use of figure captions. He displays admirable ingenuity in adding to these captions not only technical details about the experiments being illustrated, but also sociological comments and other interesting digressions that would have disrupted the flow of ideas in the text. In fact, as the caption to a reproduction of Maxwellian postcards shows (Fig. 4.5), one can even use this format to give brief explanations of important equations. As a reader who sometimes enjoys an author's footnotes more than his text, I look forward to a new era of scholarly publishing in which the most fascinating parts of a book will be found underneath its illustrations. □

Stephen G. Brush is a Professor in the Department of History and the Institute for Physical Science and Technology at the University of Maryland, College Park.

Population modelling: how are large mammals different?

Hans Kruuk

Dynamics of Large Mammal Populations. Edited by Charles W. Fowler and Tim D. Smith. Pp.477. ISBN 0-471-05160-8. (Wiley: 1982.) £33.25, \$56.55.

FOR several decades population models have been generally successfully used in the rational exploitation of fish stocks. Such models focus on productivity, and each involves only a single species. In *Dynamics of Large Mammal Populations* the editors argue that large mammals are different from other animals in their population characteristics; reading through the various chapters, however, it appears that often the approach to population research is the same as in fisheries. Many of the species discussed are being "harvested", and population models are directed towards maximum sustainable yield, with many of the characteristics of fisheries models.

However, as pointed out by G. Caughley in this volume, single-species models may serve exploitation but they do not explain ecological relationships. It is in the relationships with their environment that large mammals differ from other organisms, as argued in the chapter by R.M. Laws (to whom the book is dedicated); as a group they show similarity in feeding strategies, social behaviour, reproduction and demographic parameters. These aspects of large mammal biology have evolved together and it is dangerous to focus on merely the last of them.

The book contains 23 chapters, many with detailed single-species population models, and several general sections on modelling and management. On the whole the contributions are slanted towards marine species, especially seals; several chapters deal with terrestrial carnivores, but the many studies on herbivores (African bovids; deer, sheep) are poorly represented. Generally, there is little cross-reference between chapters, even to the extent that two different sections on models of the Pribilof fur-seal exploitation hardly refer to each other at all. There are no summaries, which makes reading of this kind of material cumbersome.

Highlights amongst the contributions are a section "What we don't know about

the dynamics of large mammals", which details the complexity of even the simplest herbivore-plant interaction in terms of populations; there is also a good comparison of population characteristics of three species of bear, an excellent evaluation of management policies of gray seals and a clearly written decision-making framework for population management. Inevitably there are weak points, too; some of the models presented in the single-species sections make assumptions which are either hopelessly unrealistic or so general as to make the model useless and recommendations impractical. Thus, a model of wolf-moose interactions recommends an optimal moose harvest under conditions whereby wolves are either left alone or controlled only when there are few of them; even the authors were unhappy with that conclusion.

With a few notable exceptions the papers on population dynamics of specific mammals are inward-looking, single-species studies. The often-used concept of density-dependence hides possible dependence on restricted resources in the animal's environment; in the last, theoretical chapter, for instance, it is found after a long and involved argument that density-dependent changes occur in populations close to carrying capacity. Clearly, there is no true density-dependence in such cases, but dependence on, for example, populations of food species.

In all the editors have missed the opportunity to push studies of large mammal populations further along by at least attempting to integrate papers dealing with exploitation with others on resource limitations and utilization. If one is interested in managing the productivity of large mammal populations, however, this volume does provide a great deal of food for thought. □

Hans Kruuk is a Principal Scientific Officer at the Institute of Terrestrial Ecology, Banchory, Scotland.

Resolving problems in spectroscopy

Norman Sheppard

High Resolution Spectroscopy. By J. Michael Hollas. Pp.638. ISBN 0-408-10605-0. (Butterworths: 1982.) £45, \$115.

DURING the past decade the subject of molecular spectroscopy has been revitalized by the development of a wide range of new techniques. Michael Hollas's book represents a timely attempt to provide a comprehensive and balanced account of both these and the more established principles of the subject.

In the preface the author states that "In general . . . the sample is assumed to be in the gas or vapour phase"; indeed, an alternative and more accurate, if perhaps less eye-catching, title for the book could have been *Molecular Spectroscopy of Gases*. Under this rubric the exclusion of the magnetic resonance spectroscopies (little applied to gases) and the inclusion — for very good pedagogical reasons — of photoelectron spectra of gases would have been naturally assumed. Under the present title these two aspects of the book seem to be very anomalous, for the magnetic resonance spectroscopies are of very high resolution, and photoelectron spectra are much less so than others.

Naturally, a particularly interesting section of the book is concerned with those newer developments that are not, or are little, covered in existing comprehensive spectroscopy textbooks. The last two chapters — some 160 pages — are mainly devoted to such topics, including the laser and photoelectron spectroscopies, and provide a welcome and well-balanced

account of the recent advances. The earlier chapters cover more familiar ground but also provide up-to-date treatments of the subject-matter.

The book covers both experimental and theoretical/interpretational aspects of the molecular spectroscopies and has well-chosen spectral illustrations. In general the writing is well-planned and clear, careful attention has been paid to detail and there are some nice historical touches. The occasional helpful footnote shows that the author is a naturally skilled teacher.

Perfection is not to be expected in a book of such wide coverage. For example, as a vibrational spectroscopist, I note that it is a mistake to use the phrase "near-infrared" to cover all that is not "far-infrared". The normal usage of the former phrase is restricted to the high frequency overtone region — the intervening region of fundamental vibrations is nowadays termed the "mid-infrared". I was also surprised to find a comment to the effect (p.176) that the use of rotational contours in infrared vibration spectra is not a "standard" technique for identifying (the symmetries of) fundamentals — it most certainly is, despite some complications.

These are, however, minor criticisms of what is overall a very successful text. The book deserves wide usage by its intended readership of undergraduates specializing in spectroscopy, and postgraduate students and teachers of the subject. □

Norman Sheppard is a Professor in the School of Chemical Sciences at the University of East Anglia.

Pergamon journals

Incorrect dollar subscription rates were quoted for four Pergamon journals in the recent New Journals review issue of *Nature* (299, pp.507, 510, 511 and 514).

Correct prices (for 1983) are: *Insect Science and Its Application* \$85; *Chinese Astronomy and Astrophysics* \$200; *PhysicoChemical Hydrodynamics* \$120; *Bulletin of Science, Technology and Society* \$85.